Improving Aerosol Simulations:

Assessing and Improving Emissions and Secondary Organic Aerosol Formation in Air Quality Modeling

A Dissertation

Presented to

The Academic Faculty

by

Jaemeen Baek

In Partial Fulfillment of the
Requirements for the Degree
Doctor of Philosophy in the
School of Civil and Environmental Engineering

Georgia Institute of Technology

December, 2009

Improving Aerosol Simulations:

Assessing and Improving Emissions and Secondary Organic Aerosol Formation

in Air Quality Modeling

Approved by:

Dr. Armistead G. Russell
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. M. Talat Odman
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Michael H. Bergin
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Rodney J. Weber
School of Earth and Atmospheric Science
Georgia Institute of Technology

Dr. Randall L. Guensler
School of Civil and Environmental Engineering
Georgia Institute of Technology

Date Approved: August 11, 2009

To my husband and daughter for their support

ACKNOWLEDGEMENT

I am truly grateful for support and kindness from many people during my Ph.D. studies. I would like to give my sincerest appreciation to my advisor, Professor Ted Russell. He has provided invaluable guidance in my research, and been a role model of a great researcher and dedicated parents. I deeply appreciate his understanding and patience allowing me to stay home for the extended period when I had my baby. I am also thankful to my committee members, Dr. Bergin, Dr. Guensler, Dr. Odman and Dr. Weber for their valuable time and helpful discussions.

I thank my friends that I have met at Georgia Institute of Technology. First of all, I thank to Dr. Yongtao Hu for supporting me both in research and in personal concerns. Special thanks to Sunkyoung Park, Alper Unal, Sangil Lee, Di Tian, Sergey Napelenok, Dan Cohan and Amit Marmur for helping me to start my study at Georgia Tech. I also thank Burcak Kanak, Bo Yan, Gretchen Goldman, Siv Balachandran, Soonchul Kwon, Galye Hagler, Farhan Akhtar, Jorge Pachon, KJ Liao, Efthymios Tagaris, and Aika Yano who enriched my experience at the graduate school. I also thank to staffs at the school of Civil and Environmental Engineering, especially Therese Rehkopf and Andrea Be for their kindness and support. Above all, I am blessed with love of my husband, Jihoon Son, and my daughter, Kaylin. Completing the Ph.D. study would have been impossible without their support and encouragement.

Research for this dissertation was funded by the Environmental Protection Agency (grant RD 83215901 and RD83107601), and Georgia Power.

TABLE OF CONTENTS

ACKNOWLEDGEMENT iv

LIST OF TABLES vii

LIST OF FIGURES x

NOMENCLATURE xiv

SUMMARY xv

CHAPTER 1. INTRODUCTION 1

CHAPTER 2. SOURCE APPORTIONMENT OF FINE PARTICULATE MATTER IN A CHEMICAL TRANSPORT MODEL AND COMPARISON WITH RECEPTOR MODELS 7

2.1 Introduction 7

2.2 Methods 10

2. 3 Results and Discussion 16

2.4 Conclusion 57

CHAPTER 3. ASSESSING EMISSION INVENTORIES OF FINE PARTICULATE MATTER USING TRACE METALS 59

3.1 Introduction 59

3.2 Methods 61

3.3 Results and Discussion 69

3.4 Conclusion 88

CHAPTER 4. ASSESSING EMISSION INVENTORIES OF FINE PARTICULATE MATTER USING ORGANIC MOLECULAR MARKERS 89

4.1 Introduction 89

4.2 Methods 90

4.3 Results and Discussion 99

4.4 Conclusion 124

CHAPTER 5. MULTIGENERATIONAL SECONDARY ORGANIC AEROSOL 125

5.1 Introduction 125

5.2 Methods 127

5.3 Results and Discussion 136

5.4 Conclusion 154

CHAPTER 6. CONCLUSION AND FUTURE WORK 157

6.1 Major Findings 157

6.2 Recommendations for Future Research 162

APPENDIX A. IMPROVING THE POINT SOURCE EMISSIONS INVENTORY FOR GEORGIA 165

REFERENCES 187

LIST OF TABLES

Table 2.1	Source categories that are used in each of models	15
Table 2.2	Source categories in CMAQ simulation and $PM_{2.5}$ emissions in July 2001 and January 2002	17
Table 2.3	Emissions of air pollutants within Metro Atlanta area in 2002	18
Table 2.4	Performance of CMAQ simulation	21
Table 2.5	Comparison of source apportionment using a brute force method and a tracer method in CMAQ at the Jefferson Street station	23
Table 2.6	Correlation coefficients between source impacts and non-ionic $PM_{2.5}$ species observations	36
Table 2.7	Comparison of monthly averaged source impacts in CMAQ, RG, LGO, PMF and MM at Jefferson St., Atlanta	45
Table 2.8	Comparison of monthly averaged source impacts in CMAQ, CMB-RG, and PMF at Birmingham, Centreville and Yorkville	46
Table 2.9	Correlation coefficients between source impacts from different models at Jefferson St., Atlanta	50
Table 2.10	Correlation coefficients between source impacts from different models at Birmingham (AL), Centreville (GA), and Yorkville (GA)	51
Table 2.11	Total $PM_{2.5}$ concentrations of source categories treated in CMB models	53
Table 3.1	Uncertainties in trace metal simulations in CMAQ	67

Table 3.2	Averaged impacts of variance in measurements, uncertainties from CMAQ simulation and the source profile uncertainties to the total weighting factors	67
Table 3.3	Performance of CMAQ simulations using source profiles in LGO and in RG	72
Table 3.4	Values and 95% confidence interval of scaling factors (F_{RG})	79
Table 3.5	Values and 95% confidence interval of scaling factors (F_{LGO} and $F_{LGO,metal}$)	80
Table 4.1	Source categories in CMB-MM and its matching categories in CMAQ	93
Table 4.2	Organic molecular markers and its major sources used in CMB-MM	93
Table 4.3	Uncertainties in estimated organic molecular marker concentrations	97
Table 4.4	Performance of simulated organic molecular markers (July 2001)	101
Table 4.5	Performance of simulated organic molecular markers (January 2002)	102
Table 4.6	Ratio of levoglucosan and resin acids concentrations in January to July	106
Table 4.7	Correlation coefficients between measured normal alkanes in July 2001 and Janaury 2002 at Jefferson St., Atlanta.	111
Table 4.8	Scaling factors (± 95% confidence interval) for emission estimates at Jefferson St. (JST)	116
Table 5.1	Reaction rates, mass stochiometric coefficients and saturation vapor pressure of precursors of secondary organic aerosol in the updated CMAQ.	131
Table A.1	Fields that are common in more than one table	171

Table A.2	Components of the QA/QC program and functions	172
Table A.3	Results of the paper survey and the complementary survey	175
Table A.4	Error types and frequency of errors	177
Table A.5	Number of records grouped by error types	178
Table A.6	Changes of emissions due to corrections	179
Table A.7	Total emissions grouped by SCC level 1	181
Table A.8	NEI 99 point source emissions from non-EGU equipments grouped by SCC level 1	182
Table A.9	Change in emissions for 30 common companies	183
Table A.10	Increase of emissions by new companies	185

LIST OF FIGURES

Figure 1.1	Population in $PM_{2.5}$ non-attainment counties in U.S. as of December 2008	2
Figure 2.1	The domain of CMAQ modeling and SEARCH monitoring sites	13
Figure 2.2	Performance of CMAQ simulation at SEARCH sites	22
Figure 2.3	Six regions in the U.S	25
Figure 2.4	Monthly average of simulated and measured $PM_{2.5}$ species in six regions.	26
Figure 2.5	Monthly averages of source apportionment of $PM_{2.5}$ in July 2001	28
Figure 2.6	Monthly averages of source apportionment of PM2.5	29
Figure 2.7	Monthly average of source contribution of primary OC, EC and unidentified primary $PM_{2.5}$ in each region.	31
Figure 2.8	Daily source impacts of $PM_{2.5}$ at SEARCH monitoring sites in July 2001 and January 2002.	34
Figure 2.9	Daily source apportionments from CMAQ, LGO, RG and PMF at Jefferson St. in July 2001	42
Figure 2.10	Daily source apportionments from CMAQ, LGO, RG and PMF at Jefferson St. in January 2002	43
Figure 2.11	EC to OC ratios defined in speciation profiles in SMOKE, in source profiles in CMB models and identified in factor analysis in PMF	55
Figure 3.1	Source profiles that are used in RG and LGO studies	68

Figure 3.2	Comparison of measured and simulated metal concentrations. Values are daily averaged concentrations of SEARCH monitoring sites	71
Figure 3.3	Spatial distribution of a) residential fuel usage – wood, b) $PM_{2.5}$ emissions from fireplaces in each county, c) $PM_{2.5}$ emissions from woodstoves in each county in 2001	75
Figure 3.4	Scaling factors for source categories a) factors calculated using CMAQ/LGO results, b) factors calculated using the set of source profiles from the LGO study with a profile of metal processes, and c) factors using CMAQ/RG	78
Figure 3.5	Contributions of wild fire, prescribed burning, wood/bark combustion, agricultural burning, fireplaces and woodstoves, and residential yard waste burning/land clearing debris to total wood combustion	81
Figure 3.6	Changes in normalized mean fractional biases (a) and errors (b) for $PM_{2.5}$, OC, EC, and trace metals at SEARCH monitoring sites	84
Figure 3.7	Source apportionment of trace metals in July 2001 and January 2002	86
Figure 3.8	Normalized changes in scaling factors when a specific species is excluded from the regression analysis	87
Figure 4.1	Daily observations and CMAQ simulation of n-alkanes in July 2001 and January 2002, at Jefferson St., Atlanta.	104
Figure 4.2	Scatter graph of ratios of measured OC vs. EC and OC with ozone	105
Figure 4.3	Observed and simulated of levoglucosan and resin acids in July 2001 and January 2002	106
Figure 4.4	Daily observations and CMAQ simulation of hopanes and steranes in July 2001 and January 2002 at Jefferson St., Atlanta.	108

Figure 4.5	Observations and CMAQ simulation of PAHs in July 2001 and January 2002 at Jefferson St., Atlanta.	109
Figure 4.6	Normalized changes in scaling factors at JST a) in July 2001using SPRogge, b) in January 2002 usingSP_{Rogge}, c) July 2001 using SP_{Lough} and d) January 2002 using $SP_{Lough.}$	113
Figure 4.7	Scaling factors for emission impacts at Jefferson St. and the other SEARCH monitoring sites	115
Figure 4.8	Changes at JST and in $Other_{SEARCH}$ of normalized mean fractional bias and error	121
Figure 4.9	Normalized mean fractional bias and error at all SEARCH monitoring sites using different CMB source profiles	122
Figure 4.10	Scatter graph of ratios of measured OC vs. EC and OC with ozone	123
Figure 5.1	The domain of CMAQ modeling and six regions in US	135
Figure 5.2	Changes in mass stochiometric coefficients and saturation vapor pressures for low-volatility SVOC and high-volatility SVOC at T=298°K and M_0 = 10 μgm^{-3}	137
Figure 5.3	Averaged hourly (a) isoprene, (b) monoterpene, and (c) sesquiterpene emissions during July 2001 and January 2002	139
Figure 5.4	Monthly average concentrations of a) monoterpene, b) total SVOCs from monoterpene reactions, c) SOA produced from monoterpene in CMAQ base case and the updated SOA module	140
Figure 5.5	Differences in secondary organic matter between ΔH156 and ΔH40 cases. Positive values indicate decreases of SOA in the ΔH40 case.	141
Figure 5.6	Aerosol yield and C_{sat} changes at different temperature when enthalphy of vaporization ΔH = 156KJ/mole and 40KJ/mole	141

Figure 5.7	Monthly average concentrations of a) isoprene, b) total SVOCs from isoprene reactions, c) SOA produced from isoprene.	143
Figure 5.8	Monthly average concentrations of a) sesquiterpenes, b) total SVOCs from sesquiterpenes reactions, c) SOA produced from sesquiterpenes.	144
Figure 5.9	Monthly average concentrations of aged aerosol. Note that the scale of January is doubled for presentation	145
Figure 5.10	Monthly average of hourly measurements of $PM_{2.5}$ (on right Y-axis), total carbon and total carbon minus black carbon, at Yorkville, Georgia.	146
Figure 5.11	Diurnal changes in aerosol yields of monoterpenes, sesquiterpenes and isoprene in the Southeastern US.	147
Figure 5.12	Simulated and measured organic carbon in July 2001 and January 2002. a) CMAQ base case, b) measurements and c) CMAQ improved	149
Figure 5.13	Monthly median of hourly simulations in the southeastern U.S. of a) $PM_{2.5}$ b) OC, and c) gas phase species	152
Figure 5.14	Monthly median of hourly $PM_{2.5}$ simulations in six regions	153
Figure A.1	The overall process of the study	167
Figure A.2	(a) GUI of the data input program (b) Report generated to check completeness of surveyed data	169
Figure A.3	(a) GUI of QA/QC program (b) a report for complementary survey that shows units with errors and suggested correction.	173
Figure A.4	Cumulative probability distribution of ratios of changes in emissions	184

NOMENCLATURE

ASACA	Assessment of Spatial Aerosol Composition in Atlanta
CEM	Continuous emission monitors
CMAQ	Community multiscale air quality
CMB	Chemical mass balance
CTM	Chemical transport model
DDM	Decoupled direct method
EC	Elemental earbon
FDDA	Four dimensional data assimilation
IMPROVE	Interagency monitoring of protected visual environment
LGO	Lipschitz global optimizer
MM	Molecular marker
NEI	National emission inventory
OC	Organic carbon
PM2.5	Particulate matter with aerodynamic diameter less than 2.5μm
PMF	Positive matrix factorization
SEARCH	Southeastern Aerosol Research and Characterization
SOA	Secondary organic aerosol
VISTAS	Visibility Improvement State and Tribal Association of the Southeast
VMT	Vehicle miles traveled

SUMMARY

Both long-term and short-term exposure to fine particulate matter ($PM_{2.5}$) has been shown to increase the rate of respiratory and cardiovascular illness, premature death, and hospital admissions from respiratory causes. It is important to understand what contributes to ambient $PM_{2.5}$ level to establish effective regulation, and air quality model can provide guidance based on the best scientific understanding available. However, $PM_{2.5}$ simulations in air quality models have often found performance less than desirable, particularly for organic carbon levels. Here, some of major shortcomings of current air quality model are addressed and improved by using CMAQ, receptor models, and regression analysis. CMAQ modeling is performed for two months (July 2001 and January 2002) for the continental U.S., and detailed analysis of source apportionment results and scaling factors are conducted in the southeastern U.S.

Detailed source apportionment of $PM_{2.5}$ is performed using the CMAQ-tracer method suggests that wood combustion (10% of total $PM_{2.5}$ in summer and 25% in winter in the Southeast), fugitive dust (7 to 10% in the Southeast), fuel combustion (10% in winter) and mobile sources (5% in both seasons) are the largest sources of $PM_{2.5}$, followed by meat cooking and industrial processes. Source impacts of $PM_{2.5}$ simulated from CMAQ and resolved in four receptor models are compared to each other at the Southeastern Aerosol Research and Characterization (SEARCH) study monitoring sites. CMAQ identified the extended number of sources (28 vs. typically 10 in receptor models) with good performances of $PM_{2.5}$ simulations with less temporal variations in source contributions than and disagreement with those from receptor models.

Discrepancies between results from CMAQ and receptor models come from biases in input data for each model and limitation in model mechanisms. Biases in CMAQ modeling is decreased by investigating emission estimates using tracer species, such as organic molecular markers and trace metals that are used in receptor models. Comparison of simulated and observed tracer species shows some consistent discrepancies, which enables us to quantify biases in emissions and improve CMAQ simulations. For example, $PM_{2.5}$ emissions from biomass burning is overestimated by 100% in January 2002, those from mobile sources are underestimated by 50% and more in both seasons. Biases in fugitive dust emissions are largest, especially in winter.

Secondary organic aerosol (SOA) is another topic that is investigated since organic carbon is one of major components of $PM_{2.5}$ in the U.S. especially in summer. CMAQ studies on organic aerosol usually underestimate organic carbon with larger than a 50% bias. Formation of aged aerosol from multigenerational semi-volatile organic carbon is added to CMAQ version 4.5, significantly improving performance of organic aerosol simulations. An increase in SOA due to aged aerosol is maximum in the south U.S., with maximum value of 8 μgm^{-3}. In the Southeast, SOA contribution is estimated as 70% of total organic carbon. Aged aerosol also decreased discrepancies between measured and simulated hourly OC concentrations.

CHAPTER 1

INTRODUCTION

1.1. CONTEXT AND MOTIVATION

Approximately one third of the US population lives in areas with unhealthy air as defined by National Ambient Air Quality Standard (NAAQS) for fine particular matter (i.e., $PM_{2.5}$ particles with diameter smaller than 2.5 μm) (EPA, 2000). As of December 2008, there were 208 $PM_{2.5}$ non-attainment counties (as defined by National Ambient Air Quality Standard: NAAQS) (Figure 1.1). Both long-term and short-term exposure to particulate matter has been found to be associated with respiratory and cardiovascular illness, hospital admissions from respiratory causes, and premature death (Dockery and Pope, 1994; Dommen et al., 2006; Peel and Tolbert, 2002; Peel et al., 2005). It is suspected that these adverse health impacts are not only due to the mass exposure, but also the type of $PM_{2.5}$ (e.g., composition, and hence source) may change its impact on human health. Mar et al. (2003) found that cardiovascular mortality in an elderly population was associated with source-specific factors (e.g., motor vehicle exhaust and burning vegetation) suggesting the source and composition of $PM_{2.5}$ are important factors to the ensuing human health impacts. For example, significant relationships between emergency department visits due to cardiovascular disease and source impacts from mobile sources and biomass burning were found in Atlanta, GA (Sarnat et al., 2008). $PM_{2.5}$ has also been related to visibility impairment, chemical deposition and change in

the solar radiation budget via scattering, absorption of sun light, and acting as cloud condensation nuclei (Seinfeld and Pandis, 2006).

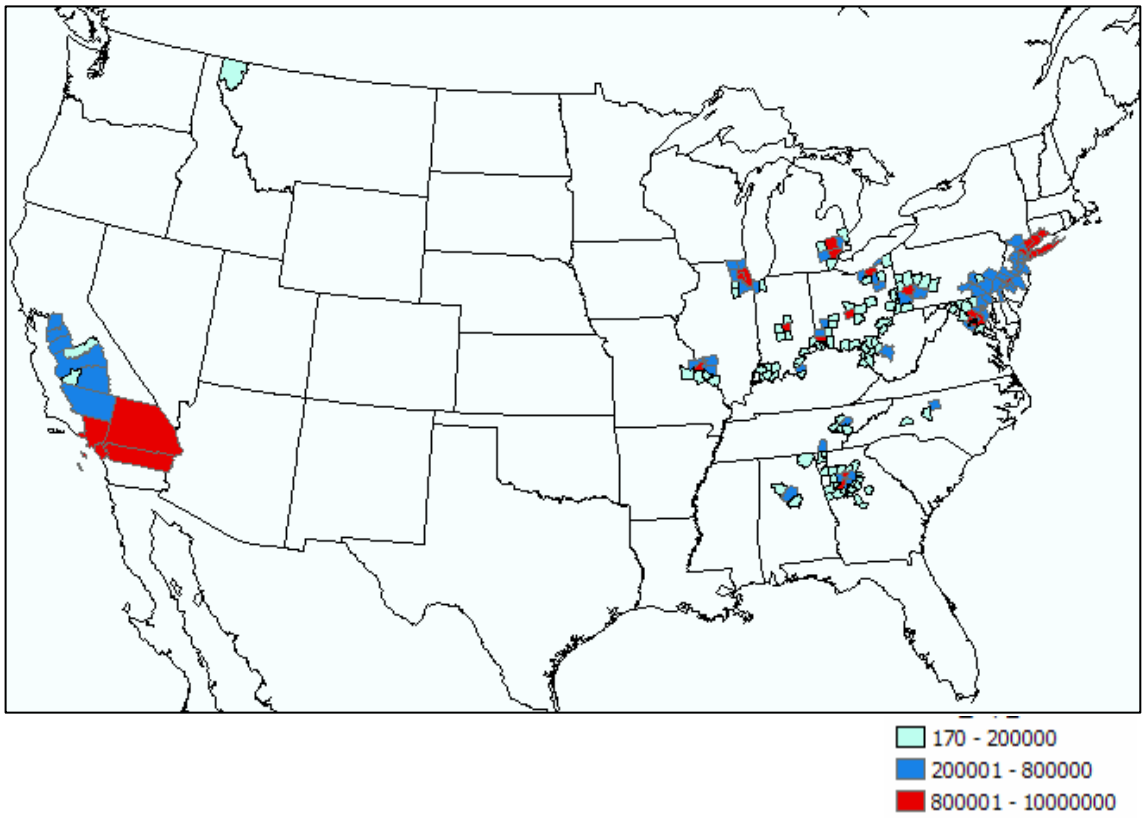


Figure 1.1 Population in $PM_{2.5}$ non-attainment counties in U.S. as of December 2008.

$PM_{2.5}$ is composed of many constituents; including sulfate, nitrate and organic carbon (OC) as major constituents. Other constituents include ammonium, elemental carbon (EC), trace metals, and usually a fraction remains un-identified materials. Composition of $PM_{2.5}$ changes by region. Overall, sulfate and organic carbon dominate in the eastern and southern U.S. and nitrate and organic carbon in the western U.S. (NARSTO, 2004). Electricity-generating units (EGU) are major sources of SO_2, which oxidize to sulfate. EGU SO_2 emissions have been decreased significantly under enforcement of Clean Air Act requirements (Evans et al., 2008). As the concentrations of sulfate decrease, reduction in ambient levels of $PM_{2.5}$ relies on identification of other major $PM_{2.5}$ constituents, such as OC and EC. Primary $PM_{2.5}$ sources, such as mobile sources, wild fires, prescribed forest burning are viewed as important sources of OC and EC. Formation of secondary organic aerosol is another major contributor to organic carbon, especially in the South and Southeast.

Our ability to more effectively manage air quality and to better identify source specific health effects relies on understanding the mass and composition of emissions from different sources (National Research Council, 2004), though both are viewed as uncertain. Air quality models are important tools for investigating sources of $PM_{2.5}$ and the impact of $PM_{2.5}$ regulations. However, performance of $PM_{2.5}$ simulations in air quality models has often been found to be less than desirable, particularly for organic carbon levels. Hence, addressing and improving some of the major shortcomings of current air quality models (particularly the Community Multiscale Air Quality [CMAQ] model) is the main objective of this thesis project. CMAQ modeling is performed over the continental U.S., with emphasis on the Southeastern U.S. using data from a variety of

sources, including the Southeastern Aerosol Research and Characterization (SEARCH) study (Edgerton and Jansen, 2004), the Visibility Improvement State and Tribal Association of the Southeast (VISTAS) study (MACTEC Inc., 2005). SEARCH is one of the U.S. EPA's super-site programs. U.S. EPA launched Super-Site programs in 1999 to characterize particulate matter, routine monitoring data that are collected as part of the health studies in large cities (EPA, 2000). As a part of the SEARCH study, there are eight monitoring stations in Alabama, Georgia, and Florida. Gas species (i.e., criteria pollutants), speciated $PM_{2.5}$, trace species have been measured since 1999. Organic molecular markers that include hazardous pollutants such as polycyclic aromatic hydrocarbons (PAHs) have been measured during certain periods as well.

1.2. SCOPE OF THIS WORK

In developing a new method for apportionment of primary $PM_{2.5}$ sources, this dissertation demonstrates how an extended number of $PM_{2.5}$ sources can be identified in CMAQ and how this information can be utilized to detect biases in both CMAQ and receptor models and to improve performance of air quality models by updating the emission inventories and detecting the level of errors. The objectives of this study are 1) to identify important sources of $PM_{2.5}$ using air quality models, 2) to improve performance of $PM_{2.5}$ simulations in CMAQ by adjusting emissions estimates, 3) to investigate reasons for biases in both CMAQ and receptor models, and 4) to advance the secondary organic aerosol mechanism in CMAQ.

This thesis is composed of five chapters following this introduction:

Chapter 2 "Source apportionment of fine particulate matter in a chemical transport model and comparison with receptor models" develops a CMAQ tracer method that apportions 28 $PM_{2.5}$ sources and implements it for the continental U.S. Source impacts analyzed by receptor models for the same period are compared with CMAQ results to identify possible biases and errors in CMAQ as well as receptor models.

Chapter 3 "Assessing emission inventories of fine particulate matter using trace metals" estimates adjustment factors for the emission inventories that indicate biases in $PM_{2.5}$ emission estimates using ridge regression analysis. Trace metals are simulated using CMAQ source apportionment results and compared with measurements in regression analysis. Applying adjustment factors to $PM_{2.5}$ emission estimates improve performance of CMAQ simulations.

Chapter 4 "Assessing emission inventories of fine particulate matter using organic molecular markers" follows the same method presented in Chapter 3, but uses organic molecular markers instead of trace metals to calculate adjustment factors for the emission inventories.

Chapter 5 "Multigenerational secondary organic aerosol" adds secondary organic aerosol formation from multi-generational semi volatile organic carbons.

Chapter 6 "Conclusion and Future Work" represents conclusions and suggests topics of further research.

Appendix contains the report on the point source survey in Georgia that demonstrates the importance of quality control of surveyed data.

(This Page Intentionally Left Blank)

CHAPTER 2

SOURCE APPORTIONMENT OF FINE PARTICULATE MATTER IN A CHEMICAL TRANSPORT MODEL AND COMPARISON WITH RECEPTOR MODELS

2.1 INTRODUCTION

Understanding sources of air pollutants is important for health studies and regulations. Many approaches have been developed to quantify the impacts of specific sources on fine particulate matter ($PM_{2.5}$: particles with an aerodynamic diameter less than 2.5μm), which is one of the criteria pollutants, that relates with adverse health impacts as well as visibility, chemical deposition and climate change. Receptor-oriented and emission-based (or chemical transport: CTM) air quality models have been proven an reliable tools for source apportionment (Carroll and Ruppert, 1988; Paatero, *et al.*, 2003; Tie, *et al.*, 2006; Watson, *et al.*, 1984).

Receptor-oriented and emission-based models are differentiated by their input data, mechanism and output data. CTMs requires emission estimates and meteorological data to simulate concentration of specific species by solving the mass balance equations in series of three dimensional grids:

$$\frac{\partial c_i}{\partial t} = -\nabla(uc_i) + R_i(c_1, c_2, ..., c_n) + E_i - S \qquad (2.1)$$

where c_i is the concentration of i at time t, u is the velocity vector, R_i is the chemical reaction term, E_i is the emission estimates of i, and S_i is the sink or removal flux of i.

Decoupled direct method (DDM-3D) sensitivity analysis is one way to to estimate source impacts in CTMs. DDM-3D computes the first-order semi-normalized sensitivity of C_i to perturbations by solving an analogous equation:

$$\frac{\partial S_{ij}}{\partial t} = -\nabla(uS_{ij}) + \nabla(K\nabla S_{i,j}) + JS_{i,j} + E_i^{'} \qquad (2\text{-}2)$$

where $S_{i,j}$ is the sensitivity of species i to parameter j, J is the ith row vector in the Jacobian matrix J, which represents the chemical interaction between species, and E_0 i is the unperturbed emission rate. The outputs of CTMs are volume-averaged, domain-wide, rather continuous source contributions.

On the other hand, receptor models rely on the observations and emission rates of tracer species from a specific sources and estimate source impacts at a specific monitoring site. The chemical mass balance (CMB) model, which is one of the receptor-based models, use ambient $PM_{2.5}$ measurements (major ions, carbon fractions and trace elements) and typical compositions of emissions from various source categories to quantify the source contribution to measured concentrations at the receptor. It is based on the following mass balance equation, which is solved for S_j (a vector of source contributions):

$$C_i = \sum_{j=1}^{n} f_{ij} S_j + e_i \qquad (2\text{-}3)$$

where C_i is the ambient concentration of chemical species i (μgm^{-3} in PM2.5), $f_{i,j}$ the fraction of species i in emissions from source j, S_j the contribution (source-strength) of source j ((μgm^{-3})), n the total number of sources, and e_i the error term. In CMB 8.0, the effective variance (EV) weighting for least squares calculations is applied, to find the best solution to the set of equations given by the equation (1).

$$\kappa^2 = \sum_{i=1}^{m} \frac{(C_i - \sum_{i=1}^{n} f_{ij} S_j)^2}{\sigma_{Ci}^2 + \sum_{i=1}^{n} \sigma_{fij}^2 S_j^2} \qquad (2\text{-}4)$$

where σ_{Ci} is one standard deviation precision of the C_i measurement; σ_{fij} is one standard deviation of the f_{ij} measurement; and m is the total number of species.

CTMs and receptor-based models often reveal qualitatively similar results, but different in detail since they use different methods and input data to apportion source contributions. Held et al. used the UCD/CIT air quality model (AQM) in the San Joaquin Valley in California, and found that the AQM and the Chemical Mass Balance (CMB) model results for a single day, by source, were correlated (slope of 0.83 with a correlation coefficient 0.55) (Held, *et al.*, 2005). Marmur et al. (2006) used the Community Multi-scale Air Quality (CMAQ) model and compared it with the CMB source apportionment results on a day-to-day basis in the Southeastern U.S. over two months. The key finding was that while both methods consistently agree on which were the major sources, they did not agree as well quantitatively, and, further, the day-to-day correlations of estimated source contributions were not as good as for longer averaging times.

A brute force method (CMAQ-BF) and the sensitivity analysis (CMAQ-DDM) have successfully been used for source apportionment of air pollutants in CMAQ. The drawback of CMAQ-BF and CMAQ-DDM is that they are computationally expensive as CMAQ-BF method requires multiple runs (i.e., Brute force method is executed by repeated runs of CMAQ with different emission inventories) and both methods need an equal number of emissions inventories to a number of source categories. Since primary aerosol is non-reactive, tracer species which can be used as fingerprints of specific

sources saving calculation time and computing resources. One example of using tracers is the CIT photochemical airshed model with tracers and has been applied to the source apportionment of fine aerosols and visibility impairment (Held, Ying, Kleeman, Schauer and Fraser, 2005;Mysliwiec and Kleeman, 2002;Schauer, *et al.*, 2001;Tie, Li, Ying, Guenther and Madronich, 2006). Here, we developed a CMAQ tracer method (CMAQ-TR) that can simulate impacts of 28 sources of primary aerosol in one model run, in order to understand the spatial and temporal changes of $PM_{2.5}$ source impacts in the U.S.

2.2 METHODS

2.2.1 CTM modeling and measurements

Emissions-based source apportionment modeling is conducted using EPA's Models-3 system, which includes the NCAR's 5th generation Mesoscale Model (MM5) version 3.5.3 (PSU/NCAR, 2003), the Sparse Matrix Operator Kernel for Emissions (SMOKE) version 2.1 (EPA, 2004), and CMAQ version 4.5 (Byun and Ching, 1999). The modeling domain covers the continental United States and parts of Mexico and Canada, with a 36km horizontal resolution and nine vertical layers (Figure 2.1). Use of a finer resolution, nested grid, did not significantly improve CMAQ results for the periods modeled (Park, 2005), and is not included here. MM5 has been used to generate the three dimensional, gridded meteorology data using four dimensional data assimilation (FDDA), and SMOKE has processed the emission inventories to create CMAQ ready input data. Base emissions are taken from EPA's NEI 2001 (EPA, 2004). Wild fires, prescribed fires and

land clearing debris combustion emissions are updated using the Visibility Improvement State and Tribal Association of the Southeast (VISTAS) 2004 results (MACTEC Inc., 2005).

Emissions from cigarette smoking and residential meat cooking are added because those sources have been shown to be significant (Schauer, *et al.*, 1996), but are not included in the NEI 2001. Emissions from cigarette smoking ($E_{cigarette}$) are calculated using county population (POP_{county}), the ratios of smokers (R_{smoker}), consumption of cigarettes ($C_{cigarettes}$) (CDC, 2006) and organic compounds emission factors ($EF_{cigarette}$) estimated by Rogge et al. (Robinson, *et al.*, 2006):

$$E_{cigarette} = POP_{county} \times R_{smoker} \times C_{cigarette} \times EF_{cigarette} \quad (2.5)$$

County level emissions are allocated to grid cells based on population distribution and assumed to have same emission rates over time. Since $PM_{2.5}$ emissions from commercial cooking processes are already included in NEI 2001 (Roelle, *et al.*, 2001), only residential meat cooking is estimated using county population, meat consumption per capita (United States Department of Agriculture, 2006), and emission factors adapted from McDonald et al. (McDonald, *et al.*, 2003). For emissions from fugitive dust, they are reduced values by applying the transport fractions (PECHAN, 2004).

Previous studies of CMB modeling with organic molecular markers (MM) suggested that vegetative detritus may be an important source of organic aerosol (Zheng, *et al.*, 2002). However, it is not included here because it is not in the NEI, and no emission factors are available. This omission will not have significant impacts on the results of this study since vegetative detritus is primarily organic carbon and has little metal content. Twenty-eight source categories are chosen to represent primary $PM_{2.5}$ sources which include 98% of the total emissions (EPA, 2004). By contrast to 28

categories, the eight to twelve source categories, such as vehicular exhaust and wood burning, typically used for source apportionment in receptor models include about 85% or less of the total primary $PM_{2.5}$ (EPA, 2003). Differences among source categories in each model is listed in Table 2.1.

CMAQ is modified to include the additional 80 tracer species tied to the primary $PM_{2.5}$ (28 OC, 27 EC and 27 other $PM_{2.5}$). Emissions of a new pseudo-species (e.g., $PM_{2.5,\text{diesel vehicle}}$, which represents primary $PM_{2.5}$ coming from diesel vehicles), is added and its ambient levels are simulated as non-reactive species having the same transport and loss properties as those of primary aerosol species calculated in CMAQ. We refer to this as the CMAQ tracer method, or CMAQ-TR. The term "tracer" in CMAQ refers to a conservative primary species, and is not specifically related with metals or organic molecular markers. The implementation of CMAQ-TR is tested by comparing the tracer-calculated impacts with brute-force assessments of the source contributions (i.e., diesel exhaust, wood burning, soil/road dust, meat cooking, and natural gas combustions). Tracer and brute force approaches agree very closely (Table 2.3).

Periods studied cover July 2001 and January 2002 when the EPA Supersites project (Solomon and Hopke, 2008) was conducting intensive measurements in the eastern US. Speciated $PM_{2.5}$ data at eight SEARCH monitoring sites is used. Four of those monitoring sites are located in urban areas (Jefferson St. (JST), Atlanta, Georgia; Birmingham (BHM), Alabama; Gulfport (GFP), Mississippi and Pensacola (PNS), Florida), three in rural area (Yorkville (YRK), Georgia; Oak Grove (OAK), Mississippi; Centreville (CTR), Alabama) and one in a suburban area (Outlying Land Field #8 (OLF), Florida) (Hansen, *et al.,* 2003). $PM_{2.5}$ species, gaseous species and trace metals are measured daily during

the studied periods. Details of SEARCH measurements are discussed elsewhere (Edgerton, *et al.*, 2005;Hansen, Edgerton, Hartsell, Jansen, Kandasamy, Hidy and Blanchard, 2003;Zheng, *et al.*, 2006).

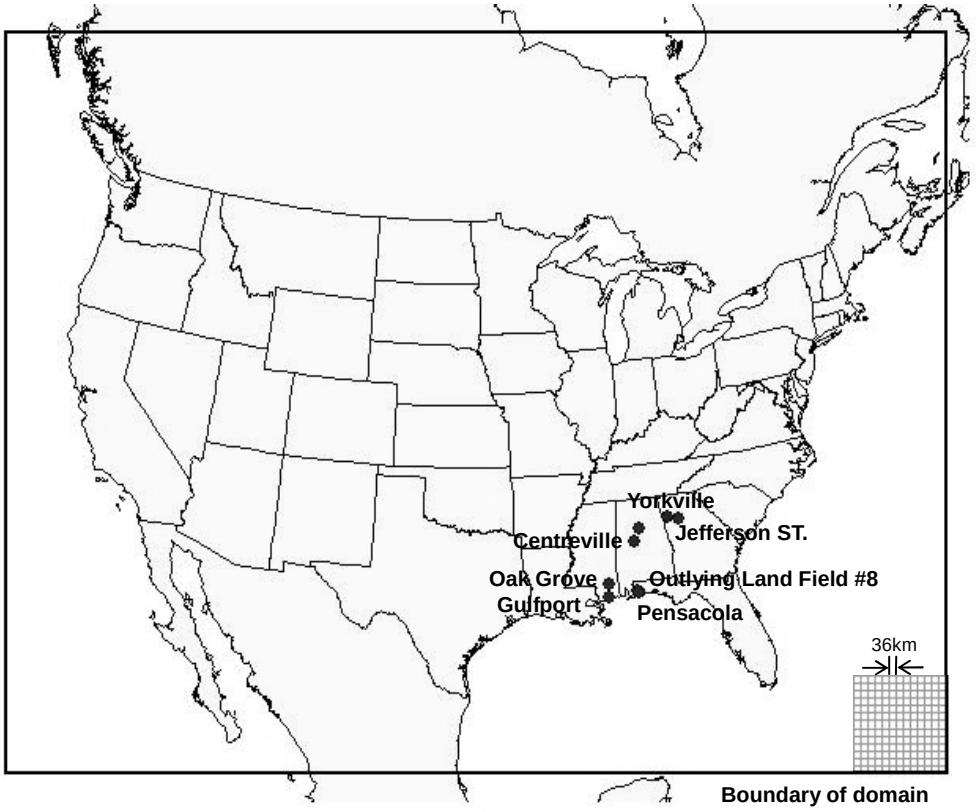


Figure 2.1 The domain of CMAQ modeling and SEARCH monitoring sites

2.2.2 Receptor models

Receptor models calculate source impacts by linear regression of source profiles (Chemical Mass Balance models, CMB) or factor analysis of ambient measurement (Positive Matrix Factorization, PMF). CMB models use ambient measurements and source profiles that are measured at emission sources (Watson, Cooper and Huntzicker, 1984). CMB applications largely rely on metal tracers (RG), though recent applications have also used organic molecular markers (MM) and gaseous species (LGO). PMF uses ambient measurements solving source profiles by factor analysis. It uses meteorological data to decide types of sources that are captured by factor analysis (Paatero and Hopke, 2002; Paatero, Hopke *et al.*, 2003). Source categories defined in receptor models are slightly different to each other and needed to be cross-matched (Table 2.1).

Table 2.1 Source categories that are used in each of models

Source types	RG	LGO	MM	CMAQ
Point sources	metal production	N/A	N/A	mineral production metal production petroleum processes paper and pulp production others
	coal combustion	coal combustion	N/A	coal combustion boilers
	N/A	N/A	N/A	Other fuel combustion
Area sources	N/A	N/A	meat cooking	meat cooking
	N/A	N/A	household heating	all natural gas boilers
Fugitive sources	soil dust	soil dust	road dust	paved road dust unpaved road dust construction dust other fugitive dust
mobile sources	mobile source	diesel engine gasoline engine	diesel engine gasoline engine	on-road diesel vehicles on-road gasoline vehicles non-road diesel engine non-road gasoline engine
wood burning	wood burning	wood burning	wood burning	wild fire prescribed burning fireplaces/woodstoves yard waste burning

2.3 RESULTS AND DISCUSSION

2.3.1 $PM_{2.5}$ emissions by source categories

Fugitive sources emit the largest amount of $PM_{2.5}$ in both seasons, followed by biomass burning, and mobile sources (Table 2.2). Emissions from fugitive sources contribute half of $PM_{2.5}$ emissions. However, most fugitive dust sources are located in the Midwest, and their contribution to $PM_{2.5}$ concentrations is not as large as to emissions elsewhere. Total $PM_{2.5}$ emitted in winter is larger than in summer by 45,000 tons, mainly due to increased wood burning in fireplaces/woodstoves, prescribed burning and other heating related fuel combustion. Emissions from most sources, such as mobile sources, fugitive sources, and industrial processes, do not have significant seasonal variations. Monthly $PM_{2.5}$ emission in metro Atlanta is shown in Table 2.3 as an example of $PM_{2.5}$ emissions in urban area.

Table 2.2 Source categories in CMAQ simulation and $PM_{2.5}$ emissions in July 2001 and January 2002 (*:reduced using NEI data using transport fractions**)**

Source categories	$PM_{2.5}$ (tons/month)	
	July 2001	January 2002
Distillate oil combustion	1,216	2,013
Wood/bark industrial combustion	667	682
Agricultural burning	5,062	11,426
Coal burning	783	2,007
Paved road dust*	10,075	10,321
Unpaved road dust*	50,583	51,899
Construction*	13,922	14,216
Other fugitive sources*	47,845	45,044
Meat-cooking	6,107	6,199
Pulp, paper and wood processing	790	226
Primary metal process	361	315
Cement Kiln	22	10
Mineral industrial processes	269	202
Petroleum and solvent evaporation	123	102
Other industrial process	536	360
Natural gas combustion – others	2,089	4,035
Natural gas combustion – residential heating	235	3,332
Diesel vehicles	6,823	7,049
Gasoline vehicles	887	2,246
Aircraft	385	409
Gasoline engine – pleasure craft	5,003	343
Diesel engine – non-road	17,832	7,530
Gasoline engine – non-road	2,729	2,644
Fireplaces/woodstoves	1,460	55,565
Residential waste & Leaf species burning	21,598	26,119
Wildfires	47,614	29,349
Prescribed burning	7,955	13,894
Others	8,245	9,723
Total	261,217	307,260

Table 2.3 Emissions of air pollutants within Metro Atlanta area in 2002 (tons/year)

Source categories	CO	NH_3	NO_X	PM_{25}	SO_2	VOC
Distillate oil combustion – external boilers	306	16	1,207	105	6,416	15
Wood/bark industrial combustion	1,811		170	33	3	50
Distillate oil combustion – internal boilers	3		26	2	1,056	0
Agricultural burning	1,255		6			193
Coal burning	8,754	252	75,901	18,259	317,588	1,116
Paved road dust				2,660		
Unpaved road dust				9,813		
Construction				3,100		
Pulp and paper, wood processing	395		5	330		531
Primary/Secondary metallurgical processes	38		45	18	0	2
Mineral industrial processes	496	182	1,943	481	2,099	170
Petroleum and solvent evaporation	17		27	0	0	3,656
Other industrial process	384	199	77	1,175	224	4,368
Meatcooking	579			1,401		203
Natural gas combustion – boilers	6420	65	9162	119	782	796
Natural gas combustion – residential	3,362	0	4,002	17	24	220
Diesel vehicles	23,655	144	76,886	2,096	1,771	4,693
Gasoline vehicles	1022,109	5,064	75,973	609	3,820	82,094
Aircraft	5,707		5,297	16	444	393
Gasoline engine – pleasure craft	5,576	0	85	36	4	2,225
Diesel engine – non-road	12,895	18	29,583	2,284	3,186	2,953
Gasoline engine – non-road	383,522	11	2,890	667	109	28,216
Fireplaces/woodstoves	18,496		231	2,554	35	6,778
Residential waste & Leaf species burning	54,205		1,766	6,849	51	5,011
Wild land fires	253	4	2	20	1	57
Prescribed burning	15,519	248	171	1,279	248	3,569
Others	18,888	12,697	5,225	1,059	1,874	98,406

2.3.2 Overall performance of CMAQ simulation

Model performance of CO, SO_2, SO_4 and ozone concentrations provides information on how well the meteorological fields and photochemical processes are simulated. CO, SO_2, and SO_4 are good pollutants to evaluate errors in meteorological fields since the emissions of CO and SO_2 are believed to be more accurately inventoried, and sulfate is a product of SO_2 oxidation. Most of the SO_2 emissions are directly measured using continuous emissions monitors (CEMs), and CO emissions, while still uncertain, are believed to be more reliable than PM emissions estimates. Normalized mean fractional bias (NMFB) and error (NMFE) are calculated using Air Quality System (AQS) (EPA, 2007) and SEARCH data (Table 2.4). NMFB for CO was -30% (AQS) and -15% (SEARCH sites only) in July 2001, and -60% (AQS) and -33% (SEARCH) in January 2002. SO_2 NMFB was -20% (AQS, Jul. 2001), -5% (AQS, Jan. 2002), -6% (SEARCH, Jul. 2001), and -20% (SEARCH, Jan. 2002). The small biases (and generally good performance) for CO and SO_2 suggest that errors in meteorological factors are relatively minor contributors to biases in PM concentrations. However, SO_2 performance changes dramatically among regions in July, while CO performance does not. In the Mountain and the Pacific, NMFBs are close to -90% in both regions, while they are between -30% to -5% in the other regions. This suggests that SO2 performance is not a good index for deciding how well meteorological factors are simulated in the mountain and the pacific U.S.

Performance of ozone with a 40ppb cutoff was -6% (AQS, Jul. 2001), -1% (AQS, Jan. 2002), 27% (SEARCH, Jul. 2001), and 26% (SEARCH, Jan. 2002) and NMFB for

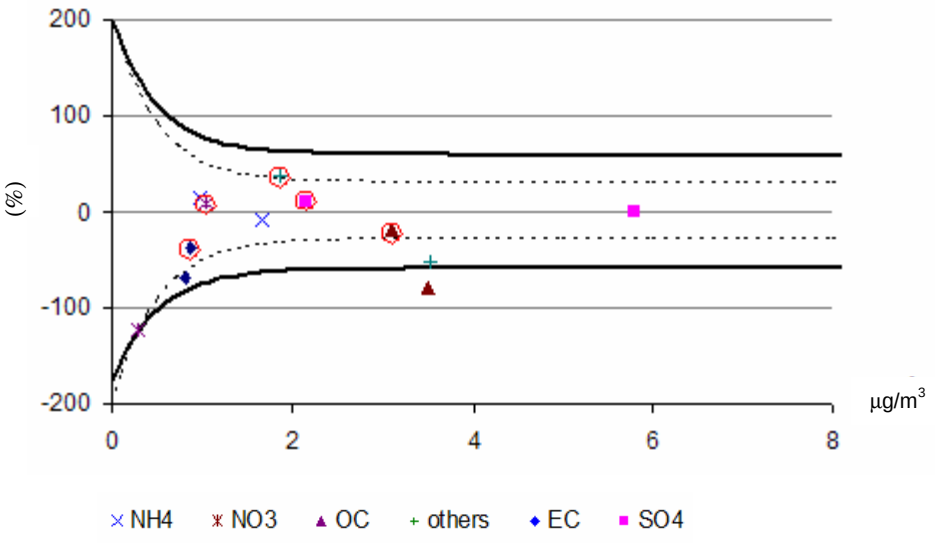
SO_4 was 7\-17% (SEARCH, Jul. 2001), and 4% (SEARCH, Jan. 2002). This indicates that the photochemical processes are also well captured, along with the related precursor emissions.

Simulation of speciated $PM_{2.5}$ was compared with the Speciated Trends Network (STN), Interagency Monitoring of Protected Visual Environments (IMPROVE) (IMPROVE, 1995), Assessment of Spatial Aerosol Composition in Atlanta (ASACA) (Butler, *et al.*, 2003), and SEARCH monitoring data (Table 2.4) (Edgerton and Jansen, 2004;IMPROVE, 1995). Based on the MFE and MFB goals, as suggested by Boylan et al. (Boylan, *et al.*, 2006), all species, except organic carbon, fall into acceptable ranges (Figure 2.2). Overall, organic carbon was underestimated about 50% of observations in summer except at JST.

Table 2.4 Performance of CMAQ simulation. Performances at STN and IMPROVE sites were calculated separately

Pollutant	No. of Records	MFB(%)	MFE(%)	Pollutant	No. of Records	MFB(%)	MFE(%)
AQS, July 2001				*Jan.2002*			
Ozone	249162	-6	26	Ozone	22564	-1	23
CO	230887	-27	75	CO	253489	-57	81
SO_2	301750	-20	83	SO_2	310145	-5	79
$PM_{2.5,}$ July 2001				*Jan. 2002*			
EC	1800	-17	67	EC	1870	22	76
NH_4	1760	-28	59	NH_4	1890	5	67
NO_3	1570	-134	145	NO_3	1815	-1	90
OC	1800	-28	76	OC	1881	-12	75
SO_4	1763	-16	46	SO_4	1894	-4	58
SEARCH July 2001				*Jan.2002*			
Ozone	1850	27	44	Ozone	267	26	60
CO	4356	-15	41	CO	4689	-33	53
SO_2	4318	-6	83	SO_2	4779	-20	75
SEARCH July 2001				*Jan. 2002*			
EC	191	-50	61	EC	188	-21	58
NH_4	185	-30	38	NH_4	182	2	60
NO_3	189	-148	153	NO_3	186	-11	90
OC	191	-52	59	OC	188	-29	56
SO_4	189	-17	38	SO_4	188	4	48
Others	155	-34	75	Others	129	60	82

a) Mean fractional bias at each SEARCH site in 2001 July and 2002 January. Mark in red circle represents the January 2002 episode.



b) Mean fractional error at SEARCH sites (July 2001 and January 2002). Mark in red circle represents July 2001 episode.

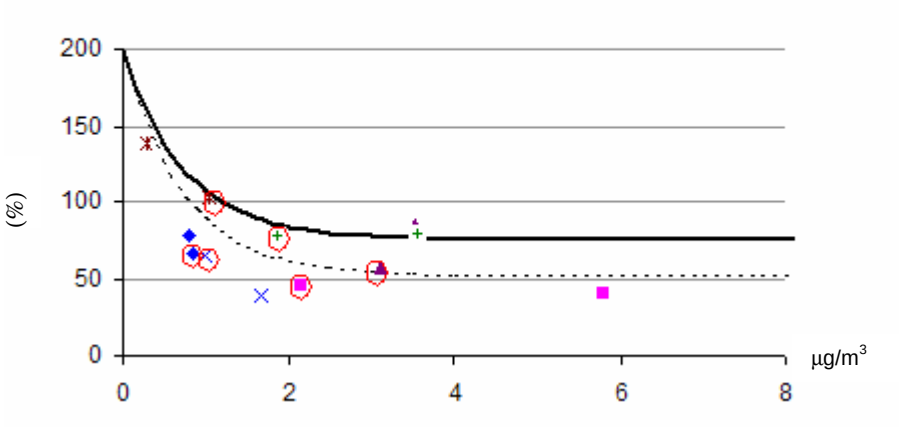


Figure 2.2 Performance of CMAQ simulation at SEARCH sites. Guide lines (solid line and dot line) are calculated based on Boylan et al.'s study (2006).

2.3.3 Comparison of CMAQ-TR with CMAQ-BF

The consistency of CMAQ-TR in estimating the impacts of sources was assessed based on the results from CMAQ-BF for July 2001 and for January 2002. Daily organic aerosol concentrations for five source categories simulated in CMAQ-TR were compared with those from CMAQ-BF at JST). MFE for five categories between CMAQ-BF and CMAQ-TR was 4.5% (July 2001) / 3.4% (January 2002) and overall mean fractional bias was 0.21% (July 2001) / 0.5% (January 2002) (Table 2.5). Slight differences between the two results were mainly due to small changes in simulation results arising from multiple runs of CMAQ.

Table 2.5 Comparison of source apportionment using a brute force method (CMAQ-BF) and a tracer method in CMAQ (CMAQ-TR) at the Jefferson Street station.

Date	STAT	Wood burning	Meat cooking	Natural gas combustion	Diesel engine exhaust	Road dust
July 2001	R^{2*}	1.00	1.00	0.99	1.00	1.00
	MFB(%)	-0.63	-2.41	-1.70	0.44	-3.83
	MFE(%)	1.45	2.73	5.22	1.71	4.15
January 2002	R^2	1.00	1.00	1.00	1.00	1.00
	MFB(%)	-0.02	0.26	1.84	2.17	-1.69
	MFE(%)	2.68	2.92	3.95	2.68	3.96

* R^2 is the correlation coefficient between simulated concentrations from two methods. Statistics are calculated at the Jefferson Street site. Negative MFB means that results from CMAQ-TR are lower than those of CMAQ-BF.

2.3.4 Source apportionment of PM2.5

Source impacts of $PM_{2.5}$ in six U.S. regions are analyzed to understand regional differences (Figure 2.3). Ionic species comprise over 50% of $PM_{2.5}$ in both seasons (Figure 2.4). CMAQ tends to simulate sulfate lower than measurements in winter. Sulfate simulation biases low in the Pacific and the Mountain regions, even in summer. MFBs of CO simulation in the Pacific (-37%) and in the Mountain (-60%) are not different from the other regions (-25% to -40%), suggesting the underestimated emissions of SO_2 is the main reason for biased low SO_4 simulations. In July, performance of SO_4 in the Mountain and in the Pacific is not as good as in the other regions (MFB -40%). There are not many SO_2 sources in both regions, and they are likely underestimated in CMAQ. Referring SO_4 performance to decide the performance of meteorological fields may not be suitable in the Mountain and the Pacific. Here, our analysis is focus in the Southeastern area, where the performance of SO_4 simulation is good (MFB 10%). Organic carbon (OC) simulations are biased low in summer, mostly due to underestimated secondary organic aerosol.

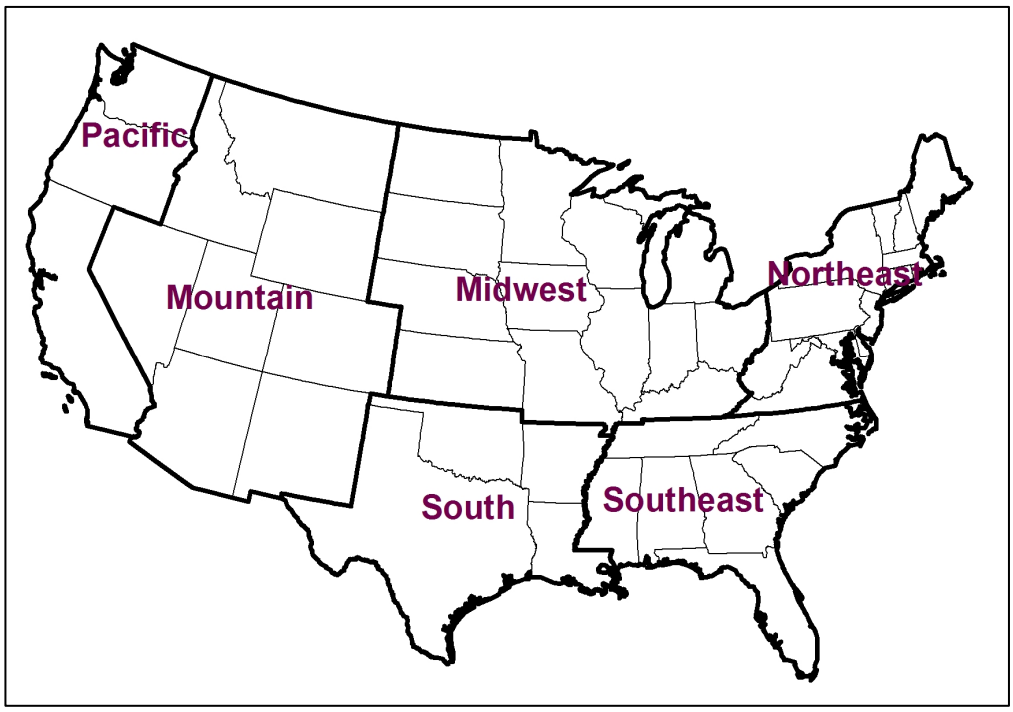
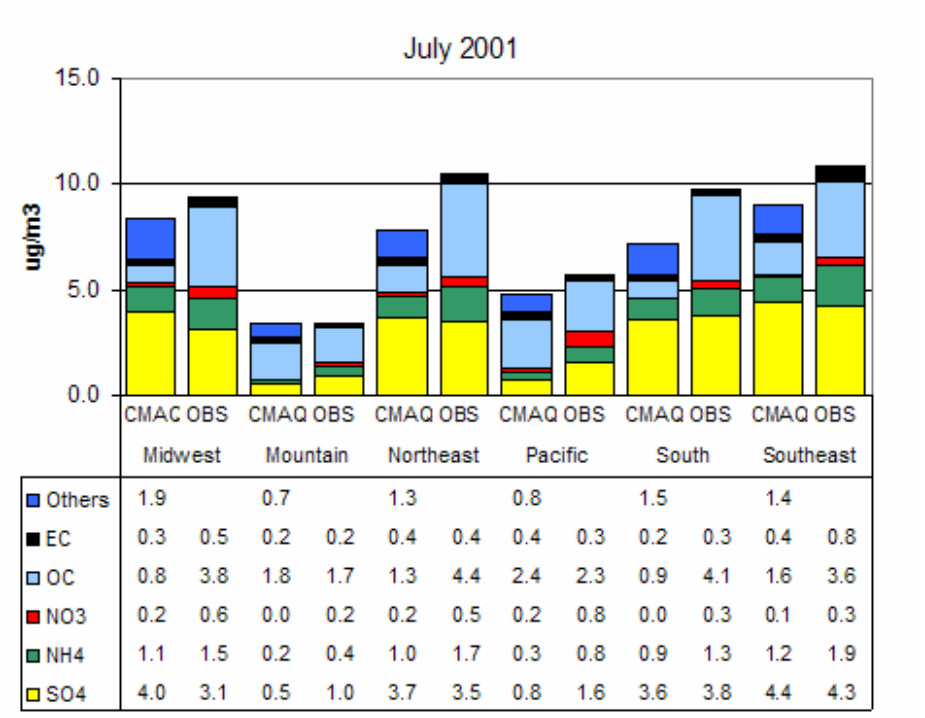


Figure 2.3 Six regions in the U.S.



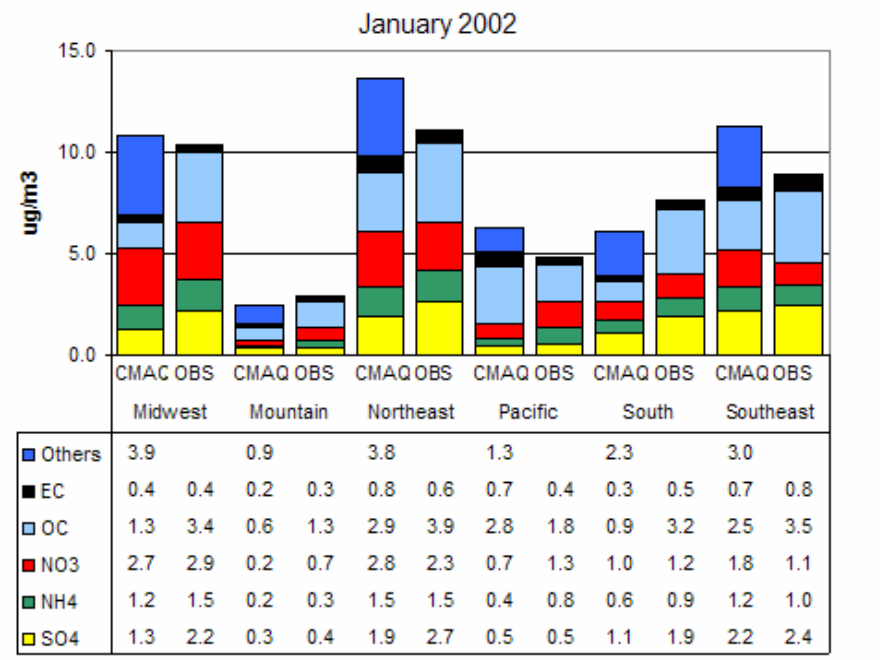


Figure 2.4 Monthly average of simulated (CMAQ) and measured (OBS) $PM_{2.5}$ species in six regions. Note that observations do not have "other" $PM_{2.5}$ composition (Others) since IMPROVE or STN do not exclusively measure it.

EC, part of OC, and other $PM_{2.5}$ originate from primary sources and their source apportionment is investigated further (Figures 2.5 and 2.6). The Midwest does not have significant $PM_{2.5}$ sources except fugitive dust from roads and other sources and wild fire is the largest source of primary $PM_{2.5}$ in the Pacific and the Northwest in July 2001. On the other hand, the Northeast, Southeast, South and part of California, where many $PM_{2.5}$ non-attainment areas are located, are influenced by a mixture of many $PM_{2.5}$ sources, including industrial processes, meat cooking, mobile sources and wood burning. Distillated oil combustion is another major source of $PM_{2.5}$ in that area.

$PM_{2.5}$ sources whose activates change with weather or are seasonally regulated according to the air pollution levels show distinct seasonality. For example, heating related fuel combustions have increased impacts in winter, as does prescribed burning. Prescribed burning increased significantly in the South, especially in Florida. $PM_{2.5}$ from construction increased in January while non-road sources do not show such an increase, suggesting that construction dust may be overestimated in January. Construction dust is closely related to non-road source activities. Emission estimates of non-road diesel engines decreased significantly in January but construction dust did not, suggesting that $PM_{2.5}$ emissions from construction dust are likely overestimated in winter.

Some sources, such as wild fires and natural gas combustions have different spatial distributions in different seasons. $PM_{2.5}$ from natural gas combustion was limited in the area where natural gas production is active. It spreads to a wider range in winter due to increased residential heating.

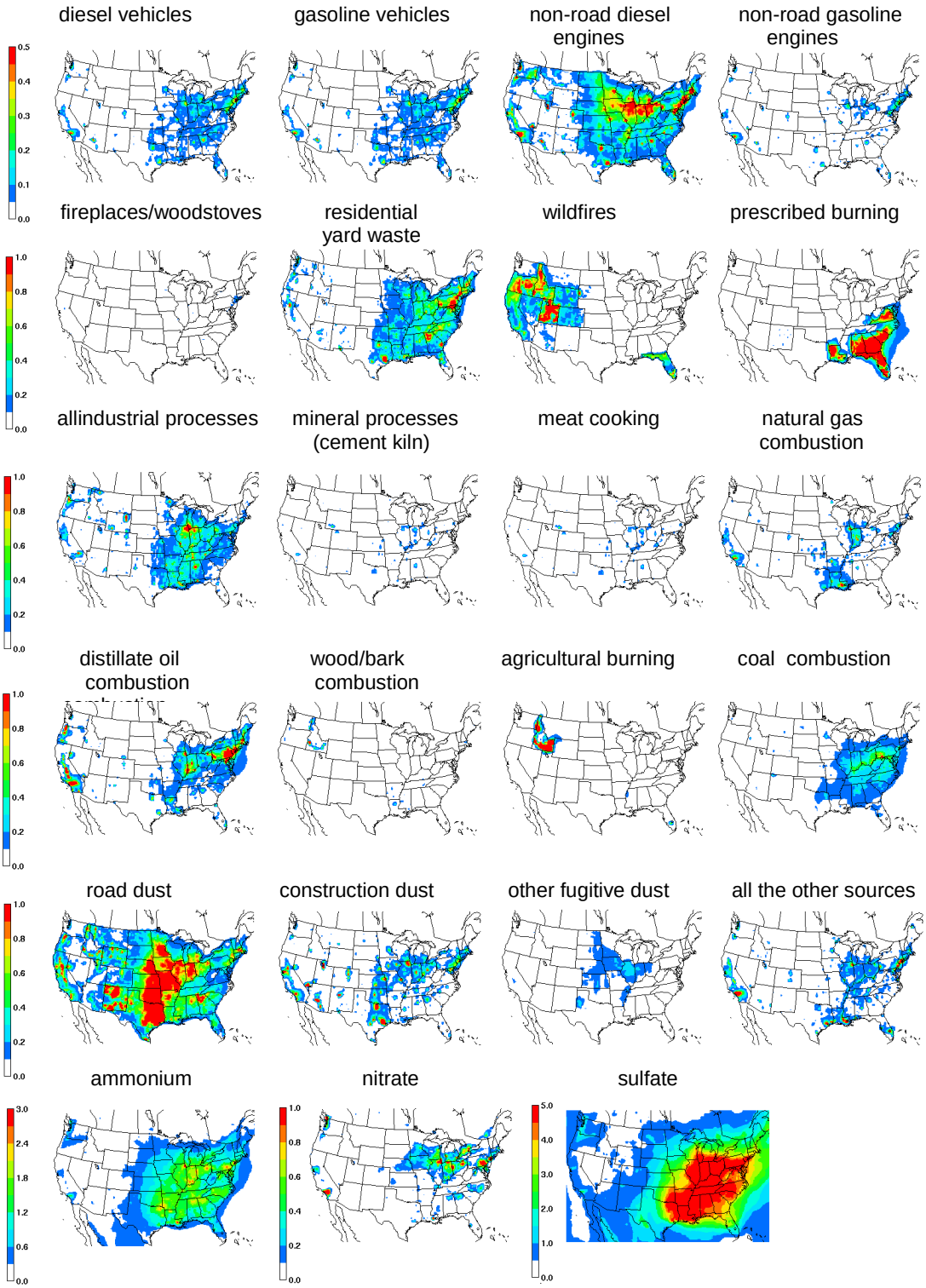


Figure 2.5 Monthly averages of source apportionments of $PM_{2.5}$ in July 2001. Units are $\mu g\ m^{-3}$

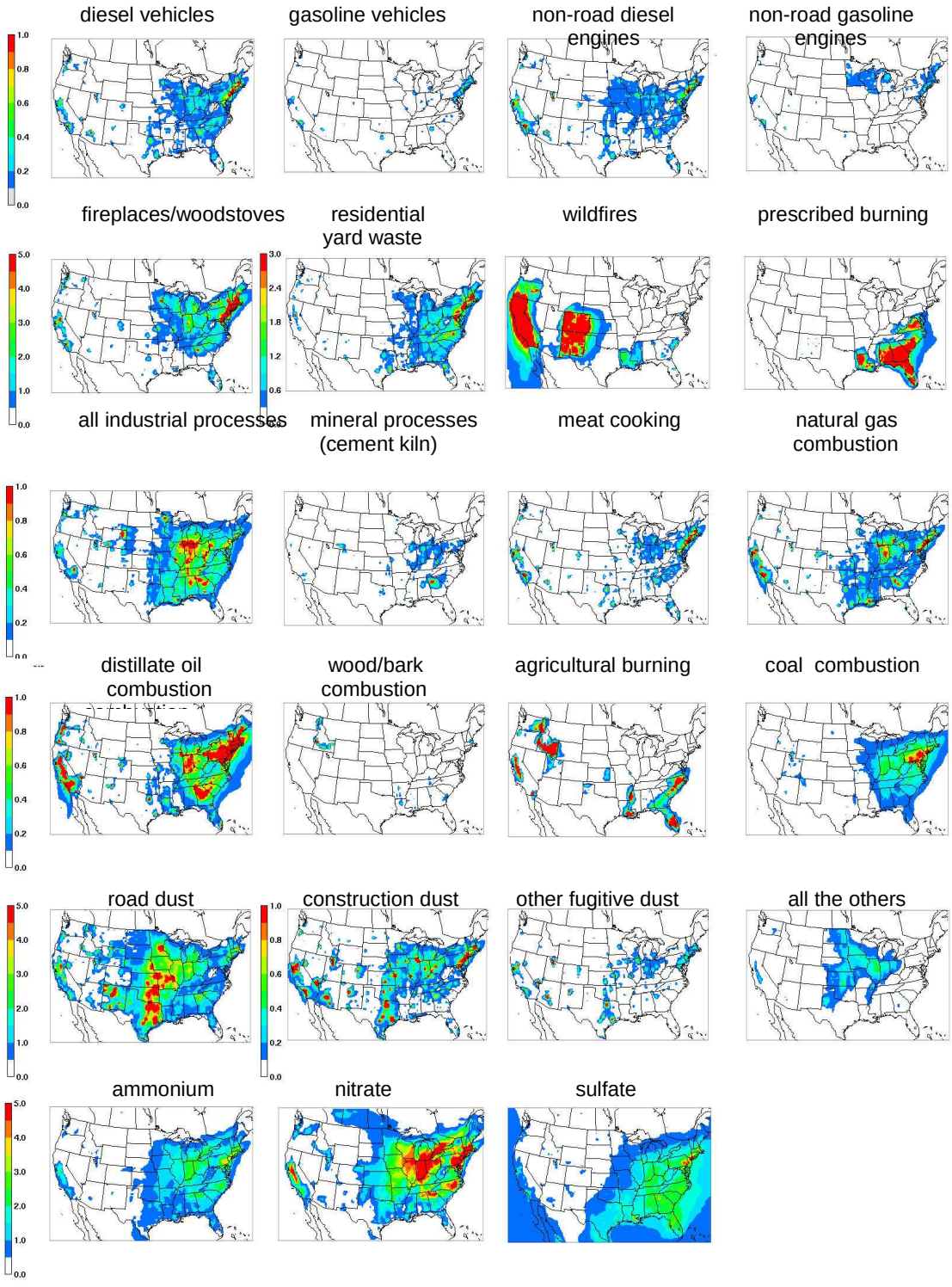


Figure 2.6 Monthly averages of source apportionments of $PM_{2.5}$ in January 2002. Units are $\mu g\ m^{-3}$

Different $PM_{2.5}$ species have different major sources (Figure 2.7). Wood combustion, diesel engine exhausts, meat cooking and industrial processes are major sources of OC. Diesel engine exhaust is the largest source of elemental carbon (EC) in January in most areas except the Pacific and the Mountain. Wood combustion becomes the dominant source of EC in winter due to increased heating and prescribed burning in the U.S except the Midwest. Other $PM_{2.5}$ mainly comes from fuel combustion, industrial processes, and fugitive dusts. It is interesting that “other” sources contribute significant amounts of EC in the South and to other $PM_{2.5}$ in the Pacific.

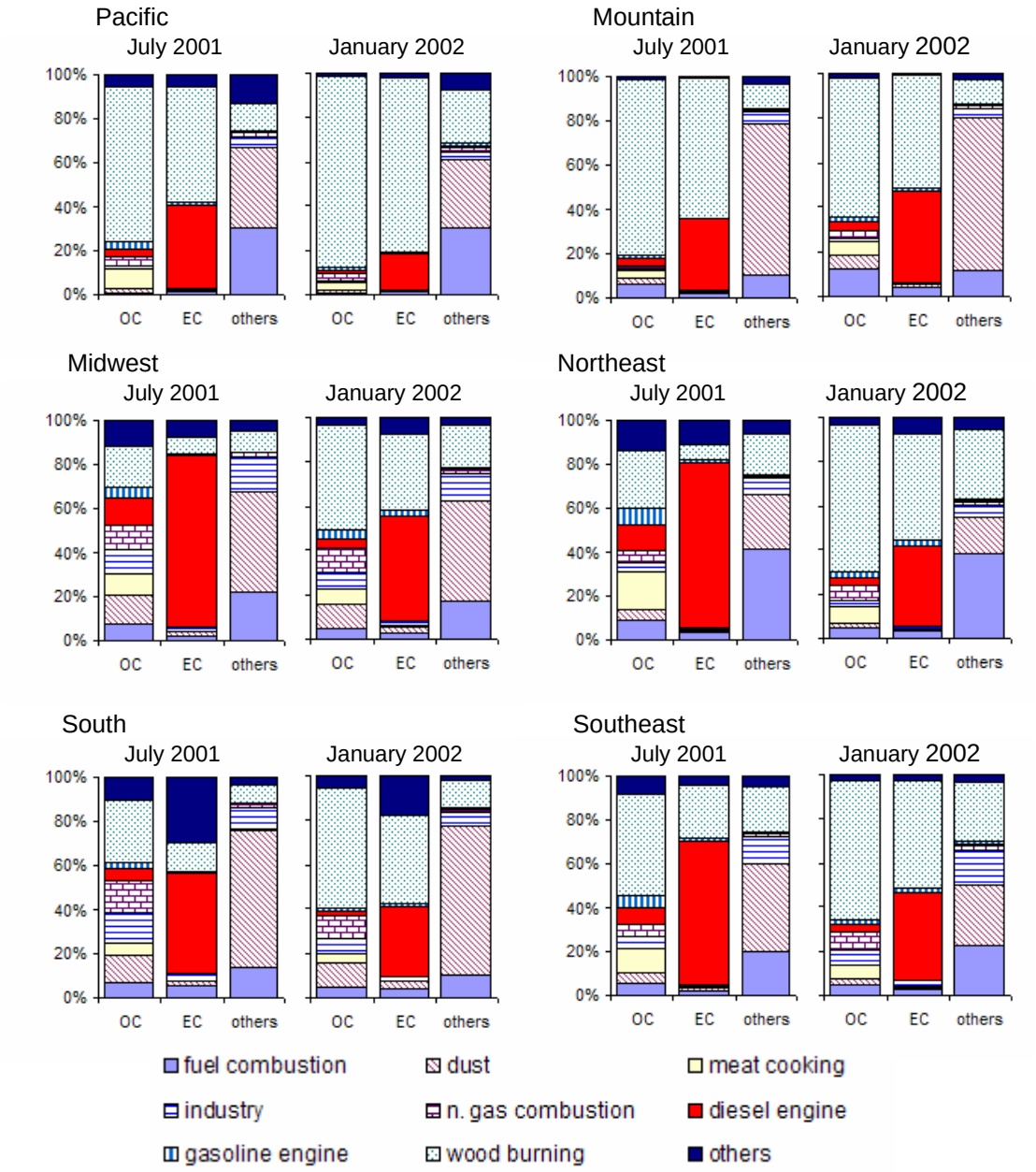


Figure 2.7 Monthly average of source contribution of primary OC, EC and unidentified primary $PM_{2.5}$ in each region.

2.3.2 Daily source apportionment of $PM_{2.5}$ at SEARCH monitoring sites

Daily source impacts of $PM_{2.5}$ at SEARCH sites are shown in Figure 2.8. Waste burning, mobile sources, and meat cooking were found to be important sources of $PM_{2.5}$ at urban sites in summer. In wintertime, fireplace/wood stove, prescribed burning, meat cooking and natural gas combustion were the biggest contributors. Due to high population density, ambient concentration of $PM_{2.5}$ from fireplaces/woodstoves at JST was three times larger than other sites except YRK. YRK is two grids away from JST in 36km domain, which is close enough to be affected by high emissions from the metro Atlanta where JST is located. For EC, non-road diesel engine exhaust, and waste burning were the largest contributors at all sites in both summer and winter. Prescribed burning was an important source of EC in January 2002 at rural sites.

The sum of daily source impacts of primary $PM_{2.5}$ and secondary organic aerosol at SEARCH sites are compared with measured EC, OC, and other $PM_{2.5}$ species (i.e., total $PM_{2.5}$ mass concentrations subtracted with sulfate, nitrate and ammonium concentrations. “Primary PM2.5” hereafter). CMAQ simulations capture the peaks in observations very well, but biased low in summer and high in winter. The correlation coefficients among eight source categories and measured primary $PM_{2.5}$ are estimated in order to investigate the possible relationship between sources and biases in CMAQ simulations (Table 2.6). It is not surprising that all nine variables are non-negligibly correlated to each other ($R >= 0.5$) since all of them, except secondary organic aerosol, are primary and primary pollutant concentrations proportionally change with dilution of air which is mainly driven by winds and the planetary boundary layer (PBL) height.

However, the contribution of wild fires and prescribed burning in July has very weak correlation with the measured primary $PM_{2.5}$ with correlation coefficients less than 0.1, suggesting that forest burning emissions have systematic biases. This is contrary to the wood burning impacts simulations in January which have similar R-values with the measurements as the other sources since forest burning emissions in January are updated with VISTAS emission inventories that are adjusted with real forest fire events.

At CTR and OLF, CMAQ fails to capture measured $PM_{2.5}$ peaks in both seasons. However, all sources except forest fire have good correlation coefficient values with measurements at both sites in July, indicating that secondary organic `aerosol may be the main reason for biases in CMAQ simulations. On the contrary, correlation coefficients between source impacts and the $PM_{2.5}$ observations are poor at the same stations in winter, suggesting that important sources of $PM_{2.5}$ are missing or its emission estimates contain large biases. CO simulations at CTR and OLF match well with CO observations (with MFB %, and MFE %), which rules out the possibility of huge biases in meteorological fields.

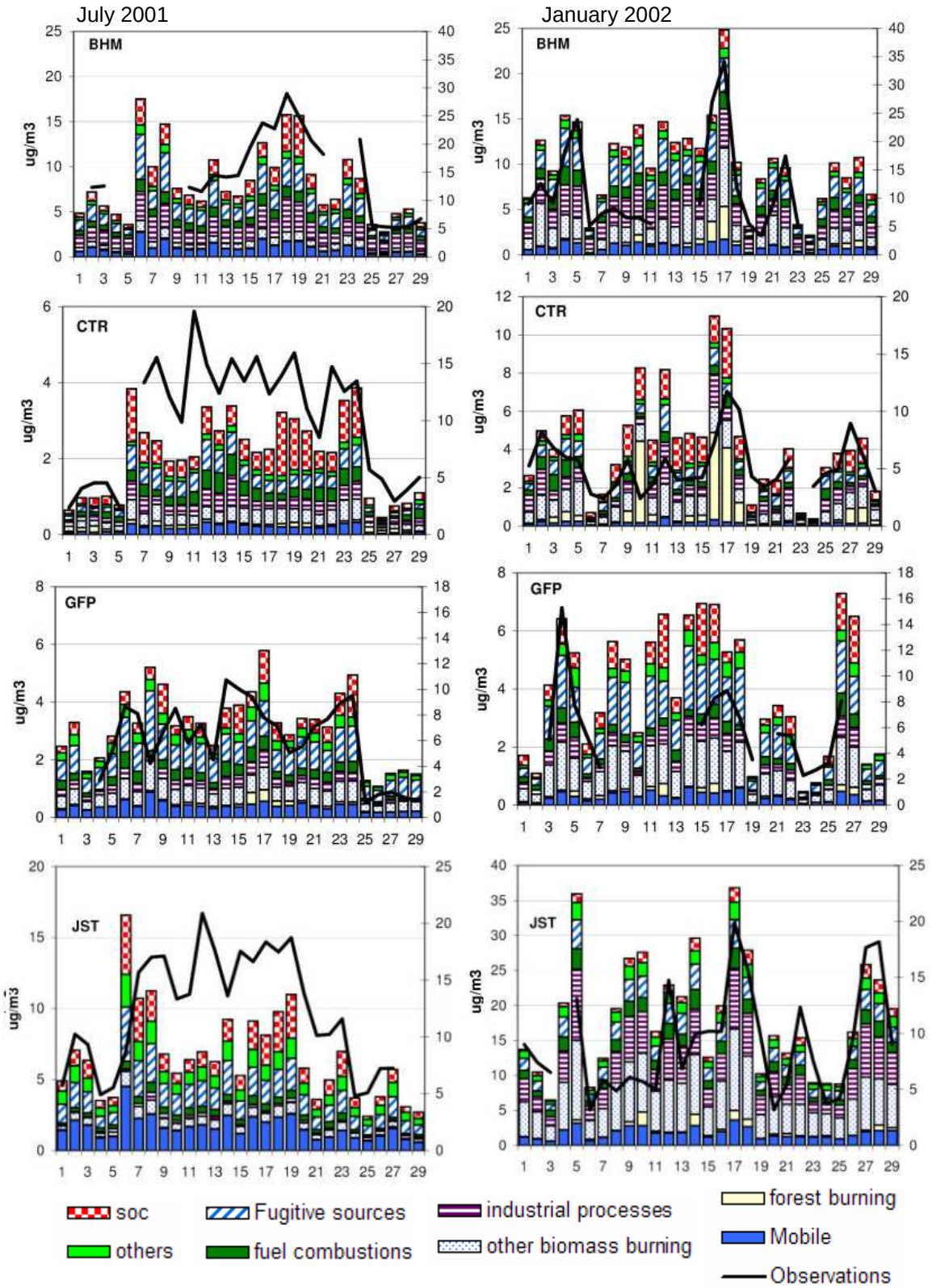


Figure 2.8 Daily source impacts of $PM_{2.5}$ at SEARCH monitoring sites in July 2001 and January 2002. Note that observation is drawn in the secondary Y axis (right)

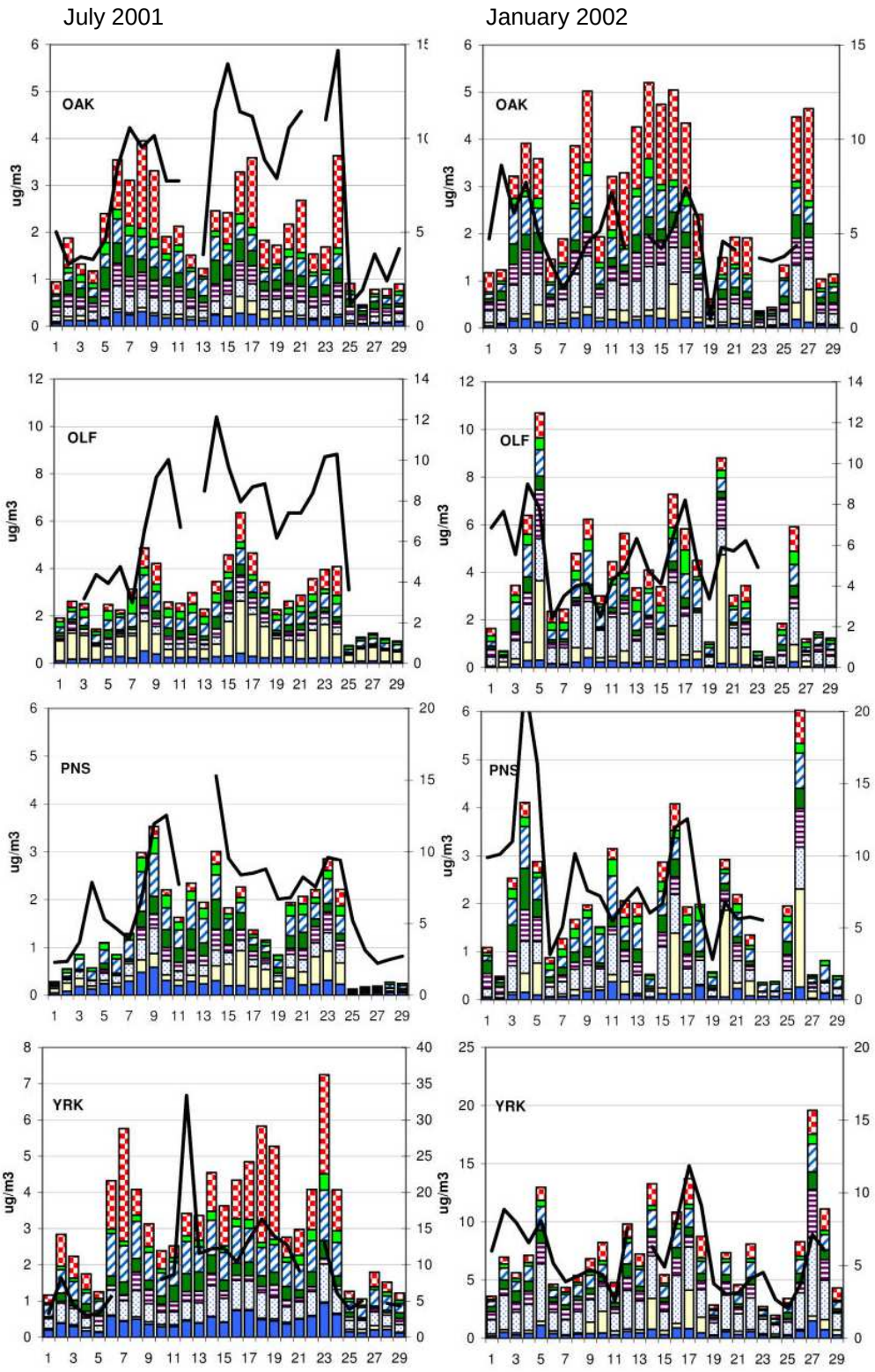


Figure 2.8. (continued)

Table 2.6 Correlation coefficients between source impacts simulated in CMAQ and observations (sum of measured EC, OC and other PM2.5)

BHM, July	total	mobile	forest	biomass	industry	fuel	fugitive	others	soc
mobile	0.96								
forest	0.16	0.13							
biomass	0.96	0.94	0.13						
industry	0.89	0.77	0.17	0.78					
fuel	0.97	0.95	0.11	0.93	0.80				
fugitive	0.98	0.95	0.13	0.95	0.81	0.98			
others	0.99	0.96	0.18	0.95	0.84	0.97	0.99		
soc	0.83	0.77	0.14	0.76	0.84	0.77	0.79	0.81	
obs	0.86	0.83	0.03	0.88	0.76	0.87	0.85	0.83	0.82
BHM, January	total	mobile	forest	biomass	industry	fuel	fugitive	others	soc
mobile	0.87								
forest	0.77	0.56							
biomass	0.87	0.58	0.70						
industry	0.83	0.72	0.49	0.61					
fuel	0.80	0.82	0.42	0.52	0.77				
fugitive	0.85	0.96	0.48	0.53	0.78	0.84			
others	0.90	0.98	0.57	0.60	0.81	0.87	0.98		
soc	0.68	0.62	0.59	0.54	0.51	0.50	0.65	0.68	
obs	0.84	0.68	0.80	0.70	0.73	0.66	0.66	0.71	0.55
CTR, July	total	mobile	forest	biomass	industry	fuel	fugitive	others	soc
mobile	0.97								
forest	-0.13	-0.16							
biomass	0.90	0.95	-0.09						
industry	0.87	0.80	-0.20	0.62					
fuel	0.93	0.90	-0.27	0.78	0.85				
fugitive	0.96	0.91	-0.07	0.87	0.76	0.84			
others	0.82	0.76	-0.25	0.75	0.65	0.71	0.79		
soc	0.61	0.64	0.16	0.72	0.35	0.42	0.67	0.48	
obs	0.81	0.74	0.01	0.75	0.62	0.76	0.83	0.68	0.63
CTR, January	total	mobile	forest	biomass	industry	fuel	fugitive	others	soc
mobile	0.81								
forest	0.82	0.39							
biomass	0.79	0.83	0.38						
industry	0.73	0.85	0.35	0.65					
fuel	0.56	0.63	0.11	0.71	0.60				
fugitive	0.75	0.94	0.38	0.77	0.70	0.52			
others	0.80	0.79	0.43	0.91	0.60	0.69	0.77		
soc	0.67	0.61	0.53	0.63	0.32	0.23	0.72	0.74	
obs	0.62	0.37	0.52	0.49	0.43	0.47	0.29	0.54	0.44

total: total primary $PM_{2.5}$ in CMAQ, mobile:mobile sources, forest: forest burning, biomass: biomass burning except forest burning, industry: industrial processes, fuel: fuel combustion, fugitive: futigive dust, soc: secondary organic carbon, obs: observations

Table 2.6 (continued)

GFP, July 2001	total	mobile	forest	biomass	industry	fuel	fugitive	others	soc
mobile	0.91								
forest	0.64	0.49							
biomass	0.88	0.86	0.43						
industry	0.87	0.68	0.69	0.60					
fuel	0.70	0.57	0.57	0.39	0.84				
fugitive	0.87	0.82	0.35	0.87	0.61	0.51			
others	0.88	0.82	0.46	0.82	0.74	0.41	0.72		
soc	0.72	0.57	0.49	0.62	0.74	0.73	0.62	0.47	
obs	0.78	0.65	0.43	0.52	0.83	0.86	0.61	0.61	0.74
GFP, January	total	mobile	forest	biomass	industry	fuel	fugitive	others	soc
mobile	0.94								
forest	0.74	0.55							
biomass	0.97	0.97	0.62						
industry	0.88	0.69	0.76	0.78					
fuel	0.57	0.33	0.66	0.39	0.76				
fugitive	0.97	0.97	0.65	0.96	0.76	0.45			
others	0.94	0.88	0.66	0.91	0.86	0.50	0.89		
soc	0.67	0.47	0.92	0.54	0.77	0.65	0.60	0.59	
obs	0.78	0.72	0.48	0.74	0.77	0.53	0.75	0.80	0.49
JST, July	total	mobile	forest	biomass	industry	fuel	fugitive	others	soc
mobile	0.93								
forest	-0.25	-0.21							
biomass	0.96	0.83	-0.32						
industry	0.44	0.15	-0.13	0.53					
fuel	0.94	0.78	-0.33	0.94	0.63				
fugitive	0.98	0.85	-0.29	0.97	0.50	0.95			
others	0.98	0.92	-0.19	0.93	0.32	0.87	0.94		
soc	0.72	0.68	-0.36	0.71	0.32	0.70	0.69	0.69	
obs	0.71	0.51	-0.27	0.77	0.65	0.81	0.75	0.62	0.47
JST, January	total	mobile	forest	biomass	industry	fuel	fugitive	others	soc
mobile	0.93								
forest	0.65	0.74							
biomass	0.98	0.96	0.65						
industry	0.89	0.69	0.33	0.81					
fuel	0.89	0.77	0.52	0.82	0.82				
fugitive	0.97	0.91	0.70	0.96	0.81	0.80			
others	0.99	0.94	0.67	0.97	0.86	0.86	0.96		
soc	0.80	0.74	0.76	0.76	0.69	0.62	0.83	0.79	
obs	0.61	0.46	0.36	0.54	0.70	0.52	0.59	0.59	0.76

Table 2.6 (continued)

OAK, July	total	mobile	forest	biomass	industry	fuel	fugitive	others	soc
mobile	0.92								
forest	0.54	0.41							
biomass	0.90	0.96	0.43						
industry	0.94	0.83	0.36	0.82					
fuel.	0.74	0.73	0.61	0.66	0.62				
fugitive	0.82	0.78	0.23	0.80	0.76	0.44			
others	0.65	0.49	0.05	0.47	0.76	0.15	0.48		
soc	0.70	0.68	0.31	0.78	0.69	0.44	0.62	0.45	
obs	0.80	0.88	0.44	0.89	0.68	0.73	0.72	0.23	0.70
OAK, January	total	mobile	forest	biomass	industry	fuel.	fugitive	others	soc
mobile	0.92								
forest	0.75	0.57							
biomass	0.96	0.97	0.62						
industry	0.95	0.83	0.62	0.89					
fuel.	0.72	0.49	0.56	0.58	0.73				
fugitive	0.93	0.95	0.57	0.94	0.87	0.50			
others	0.82	0.84	0.41	0.81	0.83	0.44	0.81		
soc	0.87	0.78	0.85	0.81	0.79	0.55	0.78	0.65	
obs	0.32	0.31	0.19	0.36	0.31	0.29	0.27	0.27	0.25
OLF, July	total	mobile	forest	biomass	industry	fuel.	fugitive	others	soc
mobile	0.92								
forest	0.83	0.67							
biomass	0.94	0.92	0.66						
industry	0.87	0.76	0.70	0.82					
fuel.	0.80	0.80	0.48	0.86	0.82				
fugitive	0.78	0.79	0.42	0.84	0.54	0.66			
others	0.39	0.49	-0.08	0.47	0.28	0.35	0.59		
soc	0.88	0.84	0.72	0.90	0.72	0.76	0.75	0.24	
obs	0.52	0.57	0.28	0.69	0.56	0.81	0.48	-0.01	0.61
OLF, January	total	mobile	forest	biomass	industry	fuel.	fugitive	others	soc
mobile	0.74								
forest	0.78	0.22							
biomass	0.81	0.93	0.33						
industry	0.75	0.24	0.84	0.28					
fuel.	0.69	0.51	0.38	0.52	0.52				
fugitive	0.73	0.99	0.20	0.91	0.21	0.51			
others	0.71	0.71	0.29	0.70	0.39	0.63	0.70		
soc	0.73	0.67	0.37	0.76	0.46	0.57	0.71	0.63	
obs	0.37	0.08	0.34	0.09	0.41	0.49	0.04	0.37	0.19

Table 2.6 (continued)

PNS, July	total	mobile	forest	biomass	industry	fuel	fugitive	others	soc
mobile	0.91								
forest	0.83	0.63							
biomass	0.91	0.88	0.69						
industry	0.75	0.54	0.82	0.46					
fuel.	0.80	0.88	0.42	0.81	0.40				
fugitive	0.72	0.78	0.29	0.79	0.21	0.80			
others	0.41	0.55	-0.12	0.48	-0.05	0.61	0.74		
soc	0.64	0.69	0.49	0.61	0.38	0.59	0.68	0.17	
obs	0.57	0.64	0.34	0.56	0.31	0.68	0.70	0.30	0.77
PNS, January	total	mobile	forest	biomass	industry	fuel	fugitive	others	soc
mobile	0.47								
forest	0.91	0.13							
biomass	0.49	0.79	0.15						
industry	0.80	0.25	0.73	0.21					
fuel.	0.55	0.47	0.32	0.47	0.42				
fugitive	0.49	0.97	0.14	0.85	0.22	0.54			
others	0.37	0.87	0.03	0.73	0.14	0.46	0.91		
soc	0.51	0.29	0.34	0.56	0.26	0.70	0.41	0.29	
obs	0.37	0.31	0.19	0.29	0.37	0.77	0.34	0.23	0.27
YRK, July	total	mobile	forest	biomass	industry	fuel	fugitive	others	soc
mobile	0.90								
forest	-0.35	-0.36							
biomass	0.95	0.90	-0.28						
industry	0.25	-0.09	-0.09	0.00					
fuel.	0.81	0.65	-0.52	0.61	0.51				
fugitive	0.95	0.76	-0.25	0.93	0.25	0.68			
others	0.87	0.92	-0.39	0.78	0.18	0.75	0.68		
soc	0.70	0.61	-0.28	0.68	0.11	0.50	0.75	0.50	
obs	0.57	0.37	-0.26	0.43	0.40	0.72	0.57	0.43	0.41
YRK, January	total	mobile	forest	biomass	industry	fuel	fugitive	others	soc
mobile	0.94								
forest	0.55	0.37							
biomass	0.95	0.99	0.34						
industry	0.84	0.87	0.14	0.88					
fuel.	0.60	0.42	0.12	0.49	0.55				
fugitive	0.92	0.86	0.56	0.88	0.66	0.42			
others	0.91	0.95	0.24	0.96	0.91	0.51	0.78		
soc	0.81	0.73	0.69	0.72	0.53	0.30	0.86	0.67	
obs	0.66	0.51	0.54	0.54	0.36	0.59	0.60	0.52	0.52

2.3.2 Comparison between receptor models and CMAQ

Comparison between receptor models and CMAQ results is conducted in two approaches. First, daily CMAQ source apportionment is compared with those estimated in receptor models at JST (Figures 2.9 and 2.10). Daily temporal variation of source impacts in CMAQ simulations is less than receptor models. Monthly averaged source impacts of $PM_{2.5}$ from CMAQ, LGO, RG, MM and PMF are compared in Tables 2.7 and 2.8.

Correlation coefficients are estimated to quantify level of differences between source apportionment results (Tables 2.9 and 2.10). All five model results are available at JST, and CMAQ, RG and PMF results at BHM, CTR, and YRK. To understand how strongly model results are correlated, the correlation coefficient (R, Pearson's product moment coefficient) is calculated for each month:

$$R = \frac{\left(E(XY) - E(X)E(Y)\right)}{\sqrt{E(X^2) - E^2(X)}\sqrt{\left(E\left(Y^2\right) - E^2(Y)\right.}} \qquad (2.6)$$

where E(A) is the average of the variable A. X and Y represent combinations of seven sources and sume of measured EC, OC and other $PM_{2.5.}$

Wood combustion At Jefferson St., CMAQ and MM showed the smallest wood combustion contribution in the summer (0.63 μgm^{-3} and 0.07 μgm^{-3}, respectively) and increased significantly in the winter (7.1 μgm^{-3} and 2.8 μgm^{-3}) (Table 2.7). Results from other receptor models slightly increased in winter but relatively smaller than MM. These are mainly due to the different seasonal changes in emissions from wood combustion in CMAQ, measured levoglucosan, and potassium (K), which are the important tracers for wood burning. Emissions from wood combustion, including wild fire, prescribed burning,

yard-waste open burning, fireplaces and woodstoves, are very low in the summer, and doubled in the winter due to increased heating and prescribed burning. Measured K increased slightly or decreased in the winter, while measured levoglucosan in the winter was larger than in the summer by a factor of 3 (OAK) to 16 (PNS). K in summer season may come from other sources, such as coal fired power plants but it is not shown in CMAQ results or in other receptor model results.

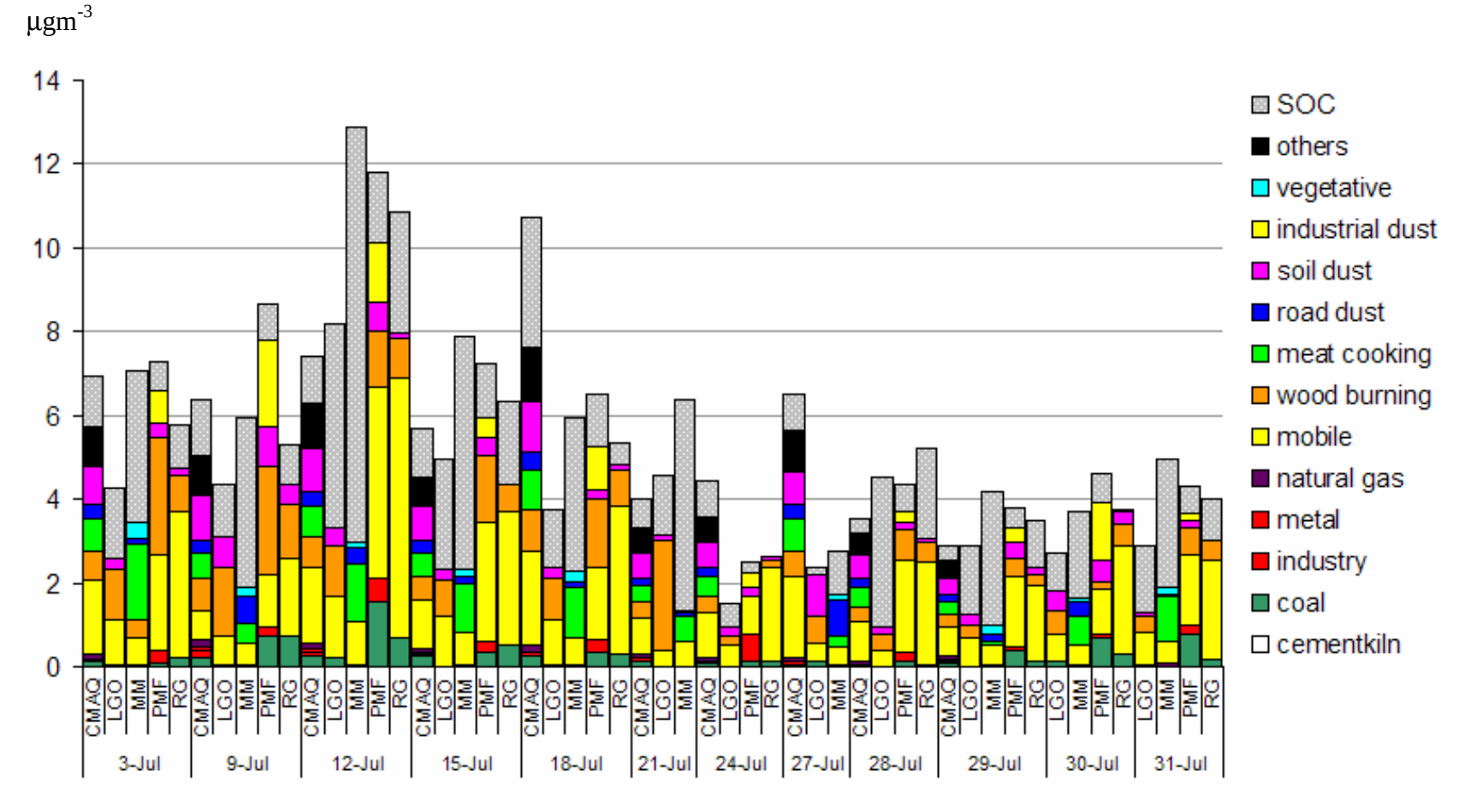


Figure 2.9 Daily source apportionments from CMAQ, LGO, RG and PMF at Jefferson St. in July 2001

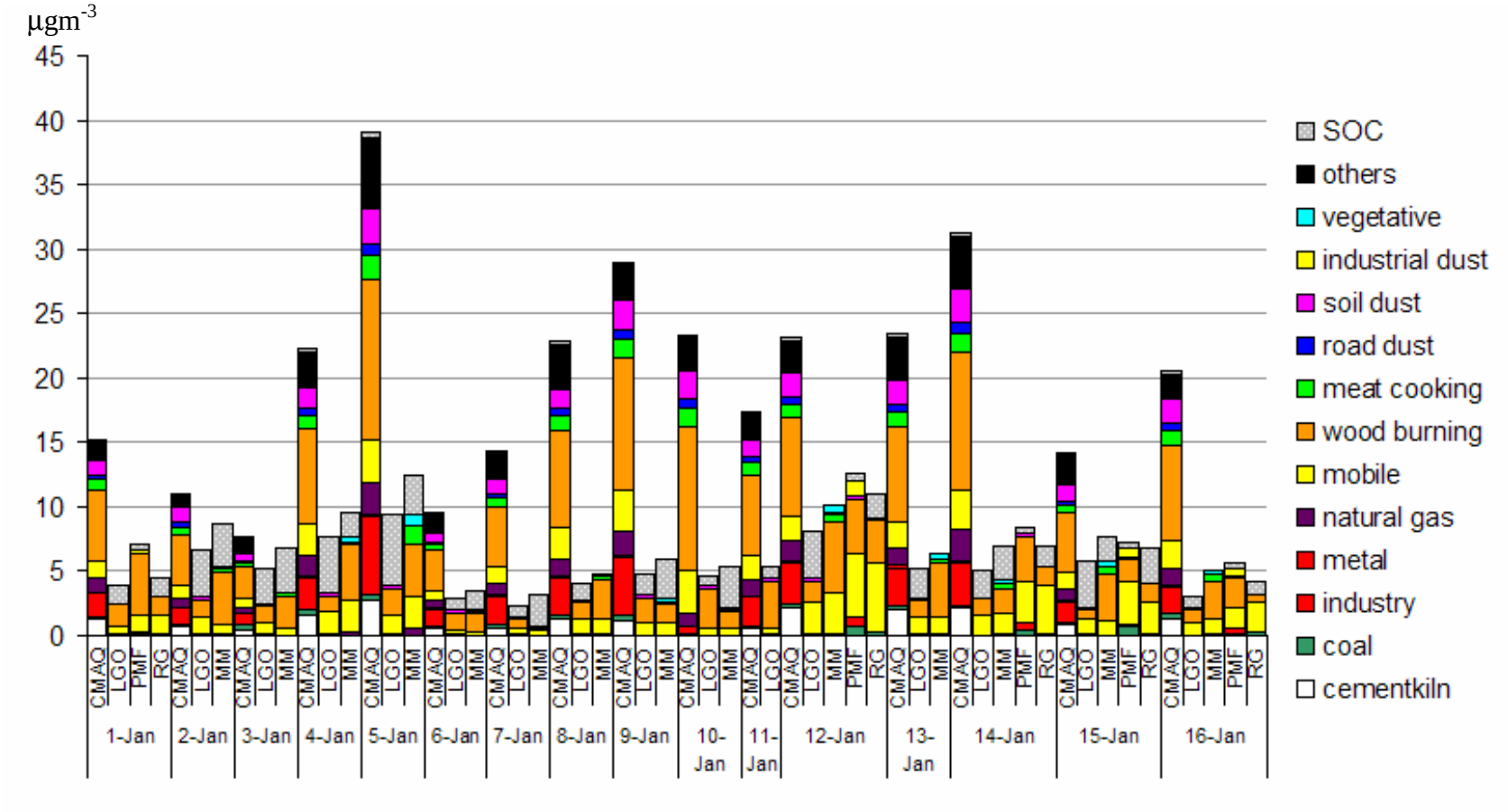


Figure 2.10 Daily source apportionments from CMAQ, LGO, RG and PMF at Jefferson St. in January 2002

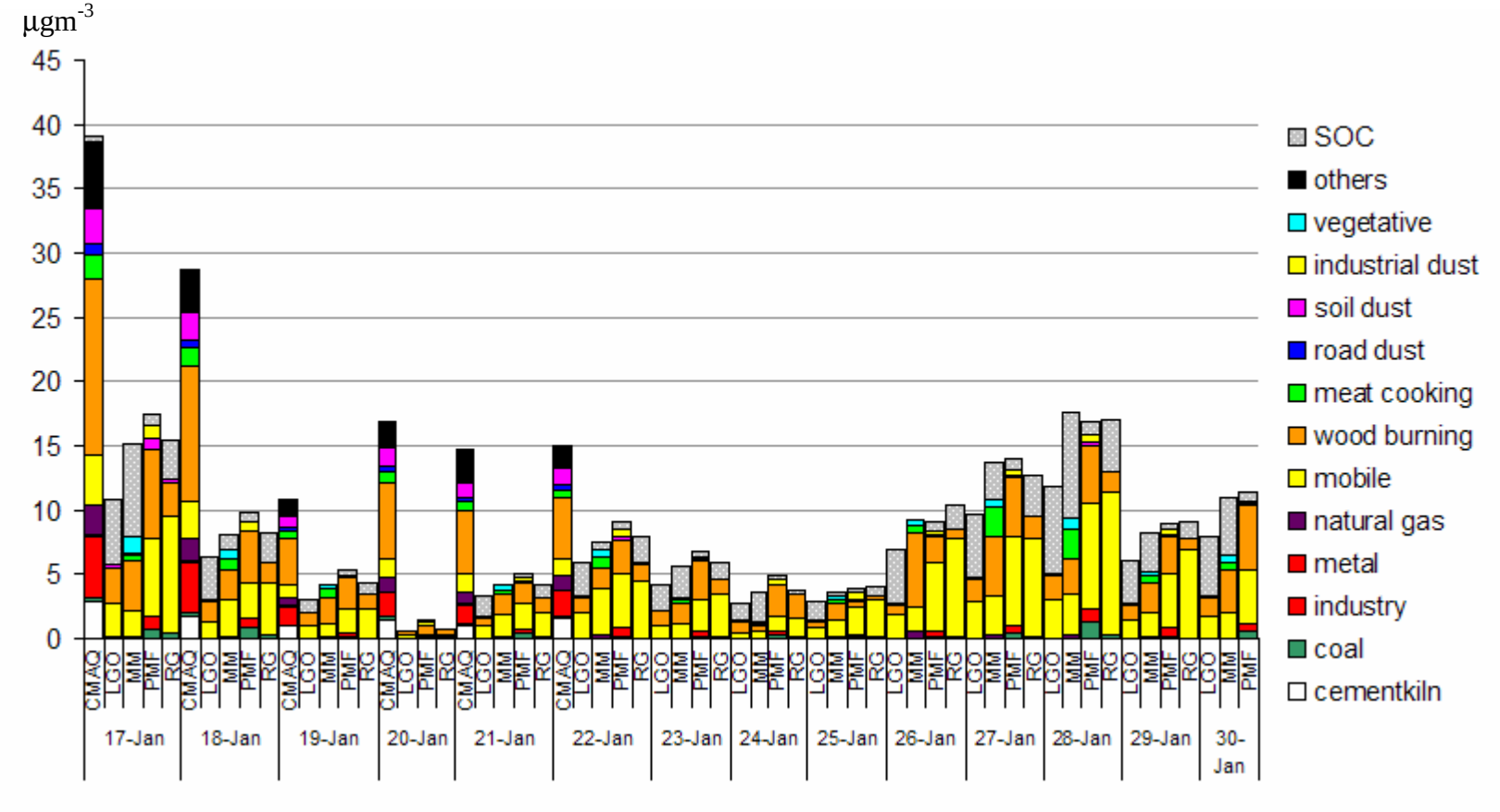


Figure 2.10 (continued)

Table 2.7 Comparison of monthly averaged source impacts in CMAQ, RG, LGO, PMF and MM at Jefferson St., Atlanta (mean ± 2 standard deviations) (units: μgm^{-3})

SOURCE	CMAQ	LGO	MM	PMF	RG
July 2001					
coal	0.19 ± 0.22	0.09 ± 0.18	-	0.69 ± 1.06	0.34 ± 0.48
industry	0.08 ± ±0.08	-	-	0.29 ± 0.44	-
metal	0.04 ± 0.08	-	-	-	-
cement kiln	0.01 ± 0.00	-	-	-	-
industrial dust	-	-	-	0.94 ± 1.40	-
natural gas	0.12 ± 0.06	-	0.06 ± 0.06	-	-
mobile	1.53 ± 1.10	0.88 ± 1.30	0.67 ± 0.98	2.20 ± 2.52	3.08 ± 3.38
diesel	1.21 ± 0.88	-	-	-	-
gasoline	0.31 ± 0.24	-	-	-	-
meat	0.67 ± 0.44	-	1.25 ± 2.08	-	-
wood	0.63 ± 0.48	0.99 ± 1.10	0.07 ± 0.36	1.04 ± 1.48	0.67 ± 0.50
soil dust	0.90 ± 0.62	0.44 ± 0.88		0.56 ± 1.08	0.24 ± 0.54
road dust	0.31 ± 0.20	-	0.43 ± 1.40	-	-
vegetative		-	0.24 ± 0.34	-	-
others	0.91 ± 0.60	-	-	-	-
SOA	1.30 ± 1.68	2.16 ± 3.04	4.52 ± 5.52	0.97 ± 0.82	1.57 ± 2.50
January 2002					
coal	0.23 ± 0.24	0.07 ± 0.14	-	0.40 ± 0.68	0.16 ± 0.20
industry	2.51 ± 2.70	-	-	0.49 ± 0.56	
metal	0.09 ± 0.12	-	-	-	-
cement kiln	1.34 ± 1.46	-	-	-	-
industrial dust		-	-	0.45 ± 0.58	
natural gas	1.31 ± 1.16	-	0.16 ± 0.32	-	-
mobile	2.01 ± 1.90	1.31 ± 1.62	1.62 ± 2.74	3.47 ± 4.14	4.30 ± 5.68
diesel	1.52 ± 1.42	-	-	-	-
gasoline	0.50 ± 0.50	-	-	-	-
meat	1.00 ± 0.88	-	0.59 ± 1.12	-	-
wood	7.06 ± 6.34	1.40 ± 1.44	2.76 ± 2.90	3.24 ± 3.24	1.41 ± 1.44
road dust	0.52 ± 0.44	-	0.01 ± 0.06	-	-
soil dust	1.60 ± 1.28	0.15 ± 0.18	-	0.16 ± 0.36	0.04 ± 0.18
vegetative		-	0.41 ± 0.56	-	-
others	2.55 ± 2.42	-	-	-	-
SOA	0.19 ± 0.18	2.64 ± 3.32	2.17 ± 4.10	0.55 ± 0.38	1.83 ± 2.26

Table 2.8 Comparison of monthly averaged source impacts in CMAQ, RG , and PMF at Birmingham, Centreville and Yorkville (μgm^{-3})

	CMAQ	PMF	RG		CMAQ	PMF	RG
Birmingham, July, 2001				*Birmingham, Jnuary 2002*			
coal	0.40±0.30	0.93±3.90	0.68±0.66	coal	0.45 ± 0.50	1.73 ± 2.74	0.56 ± 1.26
industry	0.86±0.58			industry	1.08 ± 0.86		
metal	1.02±0.72	0.43±0.50	0.87±1.52	metal	1.34 ± 0.92	0.27 ± 0.72	0.14 ± 0.26
cementkiln				cementkiln	0.01 ± 0.02		
naturalgas	0.40±0.37			naturalgas	0.75 ± 0.60		
mobile	1.02±1.00	4.01±4.68	5.61±8.16	mobile	1.20 ± 1.02	4.35 ± 7.36	7.76 ± 17.11
diesel	0.85±0.82			diesel	0.95 ± 0.80		
gasoline	0.17±0.18			gasoline	0.25 ± 0.22		
wood	0.71±0.58	0.86±0.94	0.89±0.86	wood	3.69 ± 7.28	1.18 ± 2.14	1.69 ± 3.60
meat	0.28±0.26			meat	0.36 ± 0.30		
soil dust	1.16±0.98	1.11±1.10	0.02±0.06	soil dust	1.70 ± 1..31	0.51 ± 1.08	
road dust	0.38±0.32			road dust	0.52 ± 0.40		
others	0.76±0.62			others	1.60 ± 1.20		
SOA	1.60±2.90	0.84±0.80	1.81±2.52	SOA	0.18 ± 0.20	0.83 ± 1.02	1.05 ± 1.12

Table 2.8 (Continued)

	CMAQ	PMF	RG		CMAQ	PMF	RG
Centreville, January 2002				*Centreville, January 2002*			
coal	0.18±0.24	0.66±0.90	0.19±0.48	coal	0.23 ± 0.34	0.58 ± 0.56	0.10 ± 0.20
industry	0.16±0.12	0.27±0.72		industry	0.35 ± 0.38	0.28 ± 0.66	
metal	0.09±0.20			metal	0.30 ± 0.66		
cement kiln				cement kiln			
natural gas	0.05±0.06			natural gas	0.12 ± 0.12		
Mobile	0.16±0.18		1.18±1.20	mobile	0.20 ± 0.20		1.13 ± 1.48
diesel	0.14±0.16			diesel	0.16 ± 0.16		
gasoline	0.02±0.02			gasoline	0.04 ± 0.02		
wood	0.31±0.24	5.00±4.68	0.80±0.88	wood	1.81 ± 4.16	4.07 ± 5.80	1.06 ± 1.76
meat	0.03±0.04			meat	0.06 ± 0.06		
soil dust	0.30±0.30	0.30±0.46	0.11±0.32	soil dust	0.61 ± 0.50	0.10 ± 0.20	0.07 ± 0.10
road dust	0.08±0.08			road dust	0.14 ± 0.12		
others	0.20±0.16			others	0.52 ± 0.66		
SOA	0.76±1.40	1.11±1.06	2.89±2.40	SOA	0.18 ± 0.18	0.45 ± 0.32	1.05 ± 1.16
Centreville, January 2002				*Centreville, January 2002*			
coal	0.21±0.24	1.09±1.98	0.08±0.24	coal	0.23 ± 0.28	0.79 ± 1.34	0.07 ± 0.06
industry	0.10±0.10	0.32±0.90		industry	0.66 ± 0.62	0.40 ± 0.46	
metal	0.06±0.10			metal	0.13 ± 0.18		
cement kiln				cement kiln	0.05 ± 0.14		
natural gas	0.04±0.04			natural gas	0.30 ± 0.34		
mobile	0.41±0.46		1.40±1.38	mobile	0.55 ± 0.54		1.20 ± 1.10
diesel	0.36±0.38			diesel	0.44 ± 0.44		
gasoline	0.06±0.08			gasoline	0.11 ± 0.12		
soil dust	0.86±0.64	0.20±0.60	0.04±0.20	soil dust	1.41 ± 1.10	0.13 ± 0.18	0.01 ± 0.04
road dust	0.20±0.16			road dust	0.33 ± 0.26		
wood	0.72±0.60	3.15±3.08	0.57±0.92	wood	3.51 ± 4.22	3.45 ± 3.76	0.80 ± 1.14
meat	0.08±0.14			meat	0.14 ± 0.16		
others	0.26±0.28			others	1.42 ± 1.34		
SOA	1.41 ± 2.64	1.76 ± 1.78	2.73 ± 1.98	SOA	0.16 ± 0.14	0.56 ± 0.30	1.38 ± 1.44

Road/soil dust Uncertainties of fugitive dust emissions are a well known defect in CTM. Issues such as lack of near source removal mechanism and incorrect silt content data may cause these uncertainties (Roe and Hemmer, 2004). Contribution of soil/road dust is dependent on wind fields (Marmur et al, 2006), which may be the main factor that causes errors in CMAQ simulation. At most monitoring sites, CMAQ estimated larger dust contributions than the receptor models. In all receptor model results, road dust was decreased in the winter, but increased in CMAQ simulation.

Industrial sources Different types of models have different industry categories; in CMAQ, there are six industrial categories, such as mineral processes, metal processes, cement kiln, petroleum processes, and pulp/paper processes. In the receptor models, coal-fired power plants (LGO, RG), general industrial sources high in zinc (PMF), and metal processes (RG) are used. Generally, coal combustion simulated in CMAQ is lower than the receptor models, and correlation coefficients for coal combustion between CMAQ and other receptor models are poor. There are two possible reasons for these discrepancies. The first is that the volume-averaged concentration in 36km by 36km grids in CMAQ tends to dilute source impacts of industrial sources. Another potential reason is that, coal combustion impacts are estimated high in receptor models to compensate lack of industrial sources that have non-negligible impacts on ambient $PM_{2.5}$ levels according to CMAQ source apportionments results.

Mobile sources Contributions of mobile sources estimated in CMAQ are smaller than those in receptor models at most monitoring sites. This is also related to volume-averaged concentrations in CMAQ. In the case of point sources or mobile sources, volume-averaged concentrations are non-continuously distributed and the proximity of these

sources is an important factor in determining source impacts. Receptor oriented models can capture these location specific characteristics at the monitoring station, while CMAQ dilutes emissions from the sources to the grid box. Overestimates of mobile sources in the receptor model are also suspected, considering that important sources of EC and OC such as industrial sources, meat cooking are not treated in CMB.

Meat cooking Meat cooking contributes good amount of $PM_{2.5}$ in urban area such as JST and BHM based on CMAQ results. MM returns similar amount of meat cooking contribution as CMAQ does.

Natural gas combustion Natural gas combustion in summer is very low in CMAQ and MM and is significantly elevated in winter. It increased by $1 \mu gm^{-3}$ in CMAQ, but only $0.2\ \mu gm^{-3}$ in MM at JST. Emission estimates of residential natural gas burning for heating may be biased high, since residential heating is the dominant source of natural gas combustion in winter.

`Table 2.9 Correlation coefficients between source impacts from different models at Jefferson St., Atlanta.

	Mobile Sources		Wood combustion	Coal combustion	Soil dust	Natural gas	Meat cooking	SOA
	diesel	gasoline						
July 2001								
CMAQ / RG	0.54		0.55	0.29	-0.28			-0.11
CMAQ / MM	0.54	0.23			-0.43	0.33	0.37	0.28
CMAQ / LGO	0.54	0.51	0.01	0.33	-0.31			0.30
CMAQ / PMF	0.51		0.42	0.16	-0.28			0.57
RG / PMF	0.84		0.73	0.45	0.96			0.53
RG / LGO	0.93		0.96	0.41	0.94			0.90
MM / RG	0.93							0.88
MM / LGO	0.96	0.48						0.75
MM / PMF	0.89							0.80
January 2002								
CMAQ / RG	0.75		0.93	0.83	0.48			0.53
CMAQ / MM	0.42	0.20	0.14		0.47	0.49	0.36	0.41
CMAQ / LGO	0.33	0.43	0.52	0.34	0.27			0.63
CMAQ / PMF	0.55		0.71	0.49	0.60			0.65
RG / PMF	0.95		0.75	0.66	0.94			0.90
RG / LGO	0.89		0.75	0.32	0.88			0.90
MM / RG	0.62		0.42					0.69
MM / LGO	0.89	0.55	0.14					0.65
MM / PMF	0.70		0.41					0.78

Table 2.10 Correlation coefficients between source impacts from different models at Birmingham (AL), Centreville (GA), and Yorkville (GA). The number of records is given in the parenthesis

	Mobile Sources	Wood combustion	Coal combustion	Soil dust	metal	industry	SOA
BHM, July 2001 (5)							
CMAQ / RG	0.71	0.85	-0.05		0.72		0.71
CMAQ / PMF	0.91	0.81	-0.28	0.63			0.68
PMF / RG	0.92	0.67	0.02				0.37
BHM, January 2002 (5~6)							
CMAQ / RG	0.98	0.70	0.85				
CMAQ / PMF							
PMF / RG	0.99	0.99	0.93		0.96		0.19
CTR, July 2001 (8)							
CMAQ / RG	0.93	0.41	0.67	-0.37			0.80
CMAQ / PMF	-	0.62	0.79	-0.16		0.80	0.68
PMF / RG		0.89	0.65	0.95			0.91
CTR, January 2002 (6)							
CMAQ / RG	0.68	0.94	0.50	-0.54			0.83
CMAQ / PMF	-	0.94	0.04	0.06		0.55	0.61
PMF / RG	--	0.99	0.83	0.04			0.82
YRK, July 2001 (7)							
CMAQ / RG	0.21	-0.33	0.77				0.32
CMAQ / PMF	-	-0.14	0.77	0.09		0.76	0.25
PMF / RG	-	0.84	0.60				0.94
YRK, January 2002 (7)							
CMAQ / RG	0.82	0.78	0.67				0.35
CMAQ / PMF	-	0.77	0.36	0.58		0.08	0.29
PMF / RG	-	0.93	0.39	0.74			0.59

2.3.3 Other factors that cause discrepancies between CMAQ and receptor models

(1) Sources not treated in the receptor models

In receptor-oriented models, it is assumed that source categories that are used in modeling include most important sources of $PM_{2.5.}$ This assumption is important in order to get correct source apportionment results (Zheng, Cass, Schauer and Edgerton, 2002). There are 28 source categories in CMAQ simulation, and some of them do not have matching source categories in the receptor models, i.e., there are missing source categories in the receptor-oriented models. The amount of unaccounted sources in receptor models varies from site to site, which means that conducting receptor models are suitable for some sites but not for others (Table 2.11). It is not practical to expect receptor models include all sources of $PM_{2.5}$ in real application. However, receptor models cover only half of primary PM2.5 simulated in CMAQ (40 to 70%) at SEARCH sites and uncertainties due to uncreated sources in the receptor models need to be considered when analyzing receptor model results. Quantifying unaccounted sources using CMAQ before applying receptor models at specific sites will inform of possible errors in receptor models.

Table 2.11 Total $PM_{2.5}$ concentrations (μgm^{-3}) of source categories treated in CMB models.

SITE	BHM	CTR	GFP	JST	OAK	OLF	PNS	YRK
July 2001								
MM	3.32	0.87	1.76	3.66	0.86	1.90	2.33	2.18
LGO	3.84	0.80	1.51	2.94	0.60	1.65	2.17	1.57
CMAQ	7.00	1.56	2.76	5.39	1.49	2.68	3.36	2.96
January 2002								
MM	7.37	2.81	3.06	12.73	1.76	3.18	3.96	5.98
LGO	7.60	2.68	2.65	10.25	1.27	2.72	3.63	4.69
CMAQ	12.70	4.33	4.39	20.23	2.57	4.41	5.58	8.73

CMAQ: Monthly averaged $PM_{2.5}$ concentrations in CMAQ simulations from all source categories
MM: Sum of $PM_{2.5}$ concentrations in CMAQ simulations from source categories that are treated in MM
LGO: Sum of $PM_{2.5}$ concentrations in CMAQ simulations from source categories that are treated in LGO

(2) EC to OC ratios

EC to OC ratios play an important role in source apportionments of diesel exhaust and industrial process and may result in discrepancies between source impact results in different receptor-oriented models, since they use different OC to EC ratios for the same source. In SMOKE, EC to OC ratios range from 0.1 (gasoline vehicles) to 5.1 (open burning), while RG uses one value for mobile sources and LGO uses two values: one for diesel engines and the other for gasoline engines (Figure 2.11). Contribution of diesel vehicles, gasoline vehicles to total vehicles or residential open yard waste burning to total biomass burning changes by season and time, which suggests that EC to OC ratios in the source profiles may need to be adjusted case by case. Using two different methods (TOR and TOT) to separate EC and OC may introduce additional discrepancies. EC to OC ratios defined in SMOKE are mostly measured by TOR method, but by TOT method for mobile sources. Source profiles in RG and LGO were measured by TOR while those in MM were measured with TOT. SEARCH sites uses TOR methods, so EC and OC observations are converted to TOT scales to run MM.

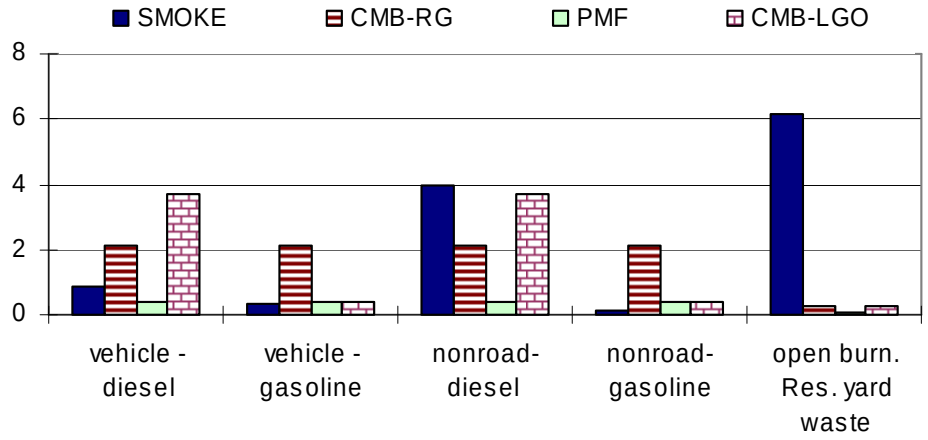


Figure 2.11 EC to OC ratios defined in speciation profiles in SMOKE, in source profiles in CMB models and identified in factor analysis in PMF.

(3) Secondary organic aerosol estimates in air quality models

CMAQ, PMF and CMB use different methods to estimate secondary organic aerosol (SOA) concentration. CMAQ uses various chemical reactions that produce secondary organic aerosol in the atmosphere using gas species and aerosol species concentrations. RG uses an EC tracer method and PMF calculates SOA from the sum of OC fraction mixed in the PMF sulfate, nitrate factor and unexplained variations. LGO assigns un-apportioned OC mass to other/secondary OC (Marmur *et al.*, 2006). Differences in methods induce differences in SOA estimates, and, in turn, source apportionment of primary $PM_{2.5.}$ For example, PMF estimates SOA lower than the EC tracer approach (Wei et al., 2005). Another study that applied PMF to estimate SOA in Hong Kong also showed the same result (Yuan et al., 2005). Receptor models are more likely to overestimate SOA. Differences between measurements and apportioned primary OC are directly related to SOC estimates, If primary OC is underestimated, SOA would be overestimated, and vice versa. Thus, the un-accounted source categories such as meat cooking in receptor models not using organic molecular markers may result in increased "other/secondary OC". On the other hand, previous studies have shown that CMAQ often underestimates secondary organic carbon significantly (Boylan *et al.*, 2006).

2.4 CONCLUSION

Ionic $PM_{2.5}$ constituents contribute more than half of $PM_{2.5}$ in the continental U.S., with sulfate being important in summer and nitrate in winter. Sulfate concentration is well simulated in most of the U.S. except the Mountain and the Pacific, where density of sources of SO_2 emissions is less than the other areas. Measurements shows that sulfate is a major component of total $PM_{2.5}$ in both regions, so It will need further investigation to adjust SO_2 emissions in the Mountain and Pacific U.S.

The other half of $PM_{2.5}$ is composed of primary $PM_{2.5}$ and secondary organic aerosol. The southern and northeastern U.S. is under the influence of variety $PM_{2.5}$ sources. Major sources of primary $PM_{2.5}$ in the Southeast are biomass burning (0.7 μgm^{-3} in summer and 7.0 μgm^{-3} in winter), mobile sources (1.5 μgm^{-3} to 2.0 μgm^{-3}), and meat cooking (1.0 μgm^{-3} in both seasons) and industrial processes (2.5 μgm^{-3} in winter). If we look into the extended source apportionment results, different sources dominate in different seasons. For example, in summer, residential yard waste burning and mobile sources are found to be important sources, while fireplace/woodstoves, prescribed burning are important in wintertime.

To assess the accuracy of simulated source contributions in CMAQ, similar results from receptor-oriented models were taken from previous studies. Source apportionment results from the five modeling approaches show general agreement, but with some discrepancies. Emission estimates of fugitive dust are likely biased, with different reasons. In summer, changes in meteorological factors that are not captured in MM5 simulation may increase or decrease fugitive dust emissions, while emission estimates in emission inventories in winter have larger biases than changes in

meteorological removal and adition mechanism. Volume-averaged concentrations of industrial processes and mobile sources in CMAQ may reduce their contribution to total $PM_{2.5}$ concentrations. Lack of important $PM_{2.5}$ sources in receptor models, such as industrial sources and meat cooking, may cause errors in source impact results.

CHAPTER 3

ASSESSING EMISSION INVENTORIES USING TRACE METALS

3.1 INTRODUCTION

Fine particulate matter ($PM_{2.5}$: particles with an aerodynamic diameter less than 2.5μm) is composed of both primary and secondary pollutants. Major ionic species of $PM_{2.5,}$ such as sulfate and nitrate, mainly come from gaseous precursor species (e.g. sulfur dioxide and nitrogen oxides) and their major sources are well known and regulated by the EPA. On the other hand, all of elemental carbon (EC), parts of organic carbon (OC) and unidentified species of $PM_{2.5}$ are directly emitted from pollutant sources. OC is from both primary and secondary pollutants, which makes accurate simulation of OC in the chemical transport models (CTMs) harder.

Many efforts to evaluate major sources of primary $PM_{2.5}$ have been made using both emission-based and receptor-based models Uncertainties in these models have been improved by updating source profiles, emission inventories, and sensitivity analysis in many studies (Gilliland, *et al.*, 2003;Lee, *et al.*, 2008).. However, the accuracy of source impact analysis in those models is still in doubt. The accuracy of estimated source contributions are further suspect due to the fact that source apportionment results from different air quality models do not agree with each other. Understanding the reasons for these disagreements and improving both types of models are crucial to better

identification of important $PM_{2.5}$ sources, and to shape more efficient air quality regulations.

In Chapter 2, differences among the Community Multi-scale Air Quality model (CMAQ), the chemical mass balance (CMB) regular and extended models are reviewed. Biases in CMAQ simulations are one of the factors that drive disagreement among models, and emission estimates are likely to be the most biased input of CMAQ modeling. To improve emission estimates, typical primary $PM_{2.5}$ species such as elemental carbon (EC), organic carbon (OC) or total $PM_{2.5}$ have been used in inverse modeling. However, using only two or three species may not provide sufficient information for inverse modeling considering the various sources of $PM_{2.5}$.

In addition to speciated $PM_{2.5,}$ we consider trace metals as additional sources of information. Trace metals, such as lead (Pb), arsenic compounds (As), manganese (Mn), chromium (Cr), nickel (Ni) and cadmium (Cd) are classified as hazardous air pollutants (HAPs) by the EPA, and are associated with adverse health impacts (Schlesinger, *et al.*, 2006). Emission rates of trace metals to particulate matter have been measured to estimate hazardous and toxic air pollutants and to be used as the input of CMB models (Holoman et al., 2005). Conducting inverse modeling by incorporating trace metal source profiles and their measurements with $PM_{2.5}$ source impact simulations gives us more source specific information than traditional $PM_{2.5}$ species (i.e., zinc represents mobile sources and metal industries, and potassium represents biomass burning and cement kilns). It also can simulate ambient trace metal levels as well as apportion sources of trace metals. In this chapter, simulated trace metal concentrations in CMAQ are compared against observations to identify biases in emission estimates, recognizing that

meteorological fields also play an important role in CMAQ modeling. The extent of meteorological input errors is addressed by examining CO, SO_x and ozone performance of CMAQ, given in Chapter 2.

3.2 METHOD

3.2.1 CTM modeling and measurements

Emissions-based source apportionment modeling is conducted using the EPA's Models-3 system, which includes the National Center for Atmospheric Research's (NCAR)'s 5th generation Mesoscale Model (MM5) version 3.5.3 (PSU/NCAR, 2003), the Sparse Matrix Operator Kernel for Emissions (SMOKE) version 2.1 (EPA, 2004), and CMAQ version 4.3 (Byun and Ching, 1999). The modeling domain covers the continental United States and parts of Mexico and Canada, with a 36km horizontal resolution. Use of a finer resolution with a nested grid did not significantly improve CMAQ results for the periods modeled (Park, 2005) and is not included here. MM5 has been used to generate three dimensional, gridded meteorological data using four dimensional data assimilation (FDDA), and SMOKE is used to process the emission inventories to create CMAQ-ready input data. Base emissions are taken from the EPA's NEI 2001(EPA, 2004). Wild fires, prescribed fires and land clearing debris combustion emissions in 2002 are updated using the Visibility Improvement State and Tribal Association of the Southeast (VISTAS) 2004 results (MACTEC Inc., 2005).

Periods studied cover July 2001 and January 2002 when the EPA Supersites project (Solomon and Hopke, 2008) was conducting intensive measurements in the eastern US.

Speciated $PM_{2.5}$ data at eight SEARCH monitoring sites is used. Four of those monitoring sites are located in urban areas (Jefferson St. [JST], Atlanta, Georgia; Birmingham [BHM], Alabama; Gulfport [GFP], Mississippi and Pensacola [PNS], Florida), three in rural areas (Yorkville [YRK], Georgia; Oak Grove [OAK], Mississippi; Centreville [CTR], Alabama) and one in a suburban area (Outlying Land Field #8 [OLF], Florida). $PM_{2.5}$ species, gaseous species and trace metals are measured daily during the studied periods. Details of SEARCH measurements are discussed elsewhere (Edgerton, *et al.*, 2005;Hansen, *et al.*, 2003;Zheng, *et al.*, 2006).

3.2.2 Ridge regression analysis

Source impacts in CMAQ-TR are combined with three sets of source profiles that were used in prior applications of CMB models in the Southeast to calculate simulated concentrations of tracer species. One set comes from Lee et al. (2008), who used a regular CMB approach (RG) which uses carbonaceous, ionic and metal species (SP_{RG}). The second set comes from Marmur et al. (2005), where the CMB-Lipschitz Global Optimizer (LGO) method was employed (SP_{LGO}) (Figure 3.1). The last set is a modified SP_{LGO} with an additional metal process profile taken from RG ($SP_{LGO\text{-}metal}$). Thus, each source has three sets of simulated metal concentrations, one from the LGO (CMAQ/LGO and $CMAQ/LGO_{metal}$), and the other from RG (CMAQ/RG) studies. The simulated concentration of metal species i (C_i) is calculated by multiplying the primary OC contribution of source j in CMAQ ($S_{OC,j}$) with the ratio of metal i to $PM_{2.5}$ in source j ($R_{i,j}$, $R_{PM2,5,j}$).

$$C_i = \sum_{j=1}^{n} S_{PM\,2.5,j} \times (R_{i,j} / R_{PM\,25,j}) \quad (3\text{-}1)$$

Species used include elemental and organic carbon, inorganic ions, aluminum (Al), silicon (Si) and other metals. Some of the 28 source categories in CMAQ-TR, such as the four categories in industrial processes and distillate oil combustion, do not have corresponding $PM_{2.5}$ source profiles in CMB studies. Thus, they are excluded from the analysis discussed below, but their emissions and species-specific impacts are simulated. Emissions from these sources are minor, comprising less than 10% of the total.

Simulated C_i's do not match with observations, and these differences imply errors in various aspects of the CTM application. Many factors, such as emission inventory errors, mis-specification of source profiles, biases in meteorological inputs, and incomplete representation of physical and chemical processes can lead to differences. Here, emission inventories are considered a major contributor to errors and, more specifically, observed biases. The importance of other factors is addressed by examining CO, SO_2 and ozone, to assess the degree to which these factors might affect comparisons of the $PM_{2.5}$ species. The magnitudes of the biases are presented as potential emissions scaling factors, which represent the amounts by which emissions must change in order to minimize weighted differences between observations and simulations. Scaling factors are estimated based on EC, OC, and ambient trace metal concentrations, assuming that they are proportional to their emissions because they are non-reactive in the atmosphere except some portion of OC. A weighted least square error fitting method is applied here (Carroll and Ruppert, 1988) with some modification. The summation of squared error (E) is calculated using measured species concentration i at site k (y_{ik}), simulated concentration of species i, from source j, at site k (c_{ijk}) and a scaling factor for source j at

site k (f_{jk}):

$$E = \sum_{ijk} (y_{ik} - c_{ijk} f_{jk})^2 w_{ik} + \lambda \sum_{jk} (f_{jk} - 1)^2 \quad (3\text{-}2)$$

The purpose of the second term in the equation (2) is to moderate changes in the emissions. λ is an adjustable length parameter that penalizes significant adjustments to source strengths. λ is adjusted to obtain physically meaningful scaling factors. In matrix form, minimizing *E* leads to an equation for the optimized scaling factors **F** (S×1 matrix) as:

$$\mathbf{F} = (\mathbf{C}^T\mathbf{W}\mathbf{C} + \lambda\mathbf{I})^{-1}(\mathbf{C}^T\mathbf{W}\mathbf{Y} + \lambda[1]) \quad (3\text{-}3)$$

where **C** is an N×S matrix of simulated trace metal concentrations (c_{ijk}), **W** is an N×N weighting factor matrix (w_{ik}), and **Y** is an N×1 vector of observed metals, OC and EC (y_{ik}). λ is chosen as the median of diagonal values of the C^TWC matrix. **I** is a S×S identity matrix and **[1]** is a S×1 column vector of ones. Fifteen **F** vectors are calculated using three sets of source profiles (CMAQ/RG, CMAQ/LGO and CMAQ/LGO_{metal} noted by F_{RG} F_{LGO}, $F_{LGOmetal,}$ respectively) and five regions (all sites, JST, urban sites, rural sites, and suburban area notated as F, F_{JST}, F_{urban}, F_{rural}, and F_{sub} respectively). $F_{RG,JST}$ represents the scaling factors at JST using CMAQ/RG data. JST is separated from the other urban sites because simulated trace metal concentrations at that site have very different characteristics. Ridge regression analysis excludes those species with data available for less than three quarters of the total measurement days. The error intervals were estimated using the variance-covariance matrix (V) of the scaling factors:

$$\mathbf{V} = (\mathbf{C}^T\mathbf{W}\mathbf{C} + \lambda\mathbf{I})^{-1}(\mathbf{C}^T\mathbf{W}\mathbf{C})\,(\mathbf{C}^T\mathbf{W}\mathbf{C} + \lambda\mathbf{I})$$

Diagonal elements of the matrix V represent the error limits in scaling factors. However, it should not be taken as accurate as uncertainties measured in experiments,

since it depends on CMAQ simulations, which are biased from true values. Here, one standard deviation given with scaling factors in order to estimate relative reliability among scaling factors (Mendoza, 2001). To estimate uncertainties related to equation (3-2) and to account for the uncertainties in the source profiles, the effective variance weighting method proposed by (Watson, *et al.*, 1984) is modified as:

$$\kappa^2 = \left[\sum_{i=1}^{n} \frac{\left(c_i^{obs} - C_i^{obs}\right)^2}{\sigma_{C_i^{obs}}^{\ 2}} + \sum_{j=1}^{l}\sum_{i=1}^{n} \frac{\left(a_{ij} - A_{ij}\right)^2}{\sigma_{aij}^{\ 2}} + \sum_{k=1}^{K} \frac{\left(c_i^{CMAQ} - C_i^{CMAQ}\right)^2}{\sigma_{C_i^{CMAQ}}^{\ 2}} + \lambda\left(C_i^{obs} - \sum_{j=1}^{J} a_{ij} SA_j^{CMAQ} \right) \right] \tag{3-4}$$

with an additional term to account for uncertainties related to CMAQ simulations. c_i^{obs} in the equation (3-3) is measured species i, C_i^{obs} is the true value of c_i^{obs}, a_{ij} is a measured ratio of species i from source j to $PM_{2.5,}$ and A_{ij} is the true value of a_{ij}. c_i^{CMAQ} is the simulated concentration of species i, and C_i^{CMAQ} is the true value of c_i^{obs} and assumed to be the same as C_i^{obs}. SA_j^{CMAQ} is the source impact of source j that is simulated in CMAQ. The uncertainties in measurements are assumed to be smaller than the other three terms, so the term $\sum_{i=1}^{n}\left(c_i^{obs} - C_i^{obs}\right)^2 / \sigma_{C_i^{obs}}^{\ 2}$ in the equation (4) is neglected in the regression analysis. Variances of measured species are added to evenly weigh each species, which have different scales of measured concentrations (e.g. EC measurements are on the order of 1 μgm^{-3} while Al and Zn are on the order of 10^2 ngm^{-3} and 10^{-2} ngm^{-3}, respectively). The weighting factor in equation (3-3) is defined as:

$$w_{ik} = \frac{1}{\mathrm{var}(y_i) + \mathrm{var}(e_{CMAQ}) + \sum_j \left(\mathrm{var}(a_{ij}) \times SA_{jk}^2\right)} \tag{3-5}$$

where var(y_i) is the variance of observed species i at all monitoring sites, var(e_{cmaq}) is the

variance of normalized CMAQ error (CMAQ simulation over observation), and var(a_{ij}) is the variance of species i in the source profile for the source category j. SA_{ijk} varies daily and w_{ik} changes accordingly.

e_{cmaq} for $PM_{2.5}$ is estimated in Chapter 3, but not for trace metals. Uncertainties in trace metals are assumed to be the same as those of $PM_{2.5}$ simulations; thus, the variance of errors in $PM_{2.5}$ simulation is converted to metal species by multiplying the ratio of average of metal measurements with the average of $PM_{2.5}$ (Table 3.1):

$$\mathrm{var}(e_{CMAQ,i}) = \left(\frac{avg(y_i)}{avg(y_{PM\,2.5}^{OBS})} \right)^2 \times \mathrm{var}(e_{CMAQ,PM\,2.5}) \qquad (3\text{-}6)$$

where avg(y_i) is the average of observed species i, $avg(y_{PM\,2.5}^{OBS})$ is the average of $PM_{2.5}$ measurements, and var($e_{CMAQ,PM\,2.5}$) is the variance of differences of simulated and measured PM2.5. var($e_{CMAQ,PM\,2.5}$) is 2 $(\mu gm^{-3})^2$ in summer and 10 $(\mu gm^{-3})^2$ in winter. $avg(y_{PM\,2.5}^{OBS})$ The contribution of each term to the total weighting factor is listed in Table 3.2.

Table 3.1 Estimated uncertainties in trace metal simulations in CMAQ (units: $[\mu gm^{-3}]^2$)

July 2001 $var(e_{cmaq,PM2.5}) = 1.86\ (\mu gm^{-3})^2$				January 2002 $var(e_{cmaq,PM2.5}) = 9.8\ (\mu gm^{-3})^2$			
species	variance	species	variance	species	variance	species	variance
Al	1.7×10^{-4}	K	3.2×10^{-4}	Al	1.1×10^{-4}	Mn	1.4×10^{-5}
Ba	4.0×10^{-5}	Mn	9.9×10^{-7}	Ba	2.0×10^{-4}	Pb	8.8×10^{-5}
Br	4.4×10^{-7}	Se	1.1×10^{-7}	Br	1.4×10^{-5}	Se	3.5×10^{-6}
Ca	1.5×10^{-4}	Si	7.6×10^{-4}	Ca	6×10^{-5}	Si	9.6×10^{-4}
Fe	2.8×10^{-4}	Zn	1.6×10^{-5}	Fe	1.5×10^{-3}	Zn	2.8×10^{-4}
				K	1.1×10^{-3}		

Table 3.2 Averaged impacts of variance in measurements (%), uncertainties from CMAQ simulation and the source profile uncertainties to the total weighting factors (July 2001)

Species	$var(e_{cmaq,metal})$	$var(C_i)$	$Var(a_{ij})xSA_j^2$
AL	0.05*	0.85	0.11
Br	0.15	0.70	0.16
Ca	0.06	0.74	0.20
EC	0.13	0.85	0.02
Fe	0.09	0.84	0.07
K	0.02	0.94	0.04
OC	0.23	0.77	0.00
Si	0.07	0.87	0.05
Zn	0.02	0.90	0.08

* Impacts range from zero (zero contribution) to 1.0 (100% perecent). Variances of measurement contribute from 70 to 90% of the weighting factors. Errors in CMAQ simulation are the largest in OC and EC, and uncertainties in source profiles are largest with Ca.

3.2.3 Source profiles

Source profiles taken from LGO, RG and Chow et al.'s studies are shown in Figure 3.1;the same wood combustion, coal combustion, and soil dust source profiles are used in LGO and RG studies (Chow, *et al.*, 2004;Lee, *et al.*, 2008;Marmur, *et al.*, 2006).

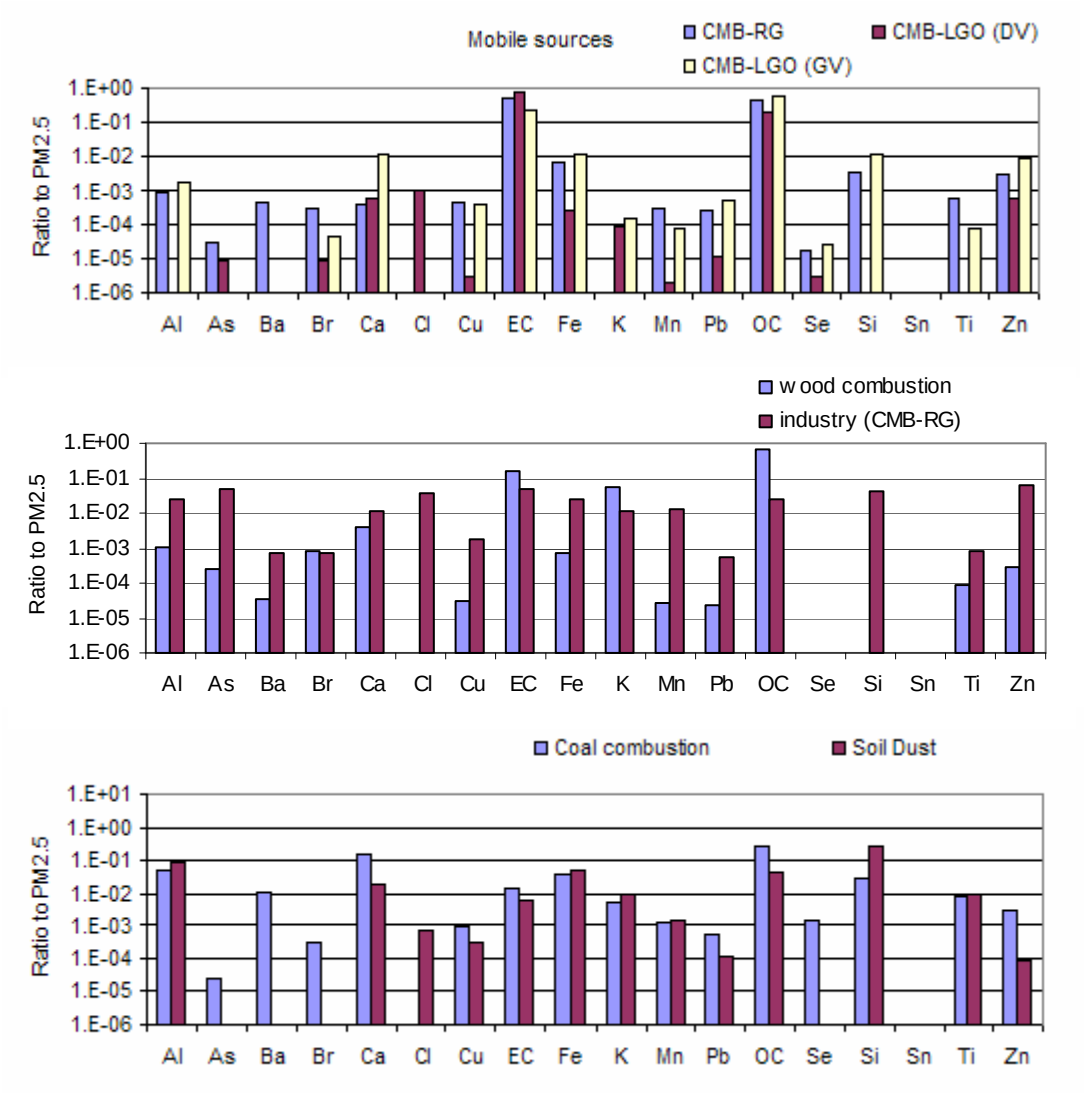


Figure 3.1 Source profiles that are used in RG and LGO studies (Lee et al., 2007; Marmur et al., 2005).

3.3 RESULTS AND DISCUSSION

3.3.2 Comparison of simulated metal species and observations

Overall, simulated and measured selenium (Se), calcium (Ca), iron (Fe), potassium (K), and manganese (Mn) concentrations were in good agreement (Figure 3.2, Table 3.3 for extensive performance statistics). Se is primarily from point sources (e.g., coal combustion sources), and Ca is abundant in industrial processes, coal combustion and soil dust emissions (Hopke, *et al.*, 2006; Marmur, *et al.*, 2005). Simulated and measured Ca were in good agreement at rural sites, and simulated Ca concentrations were higher at urban sites due to increased wood combustion, coal burning and soil dust contributions in CMAQ-TR

Potassium, which is often taken as a tracer for wood combustion, was simulated to be higher in winter than in summer. In the winter, the NMFBs of simulated K were larger at urban sites (150% at JST and BHM, and 76% at GFP) as compared to rural sites (85% at CTR and 17% at OAK). An exception is YRK, which is a rural site, where the NMFB was 140%. YRK is near Atlanta, and is often impacted by the same emissions as JST. Simulated K concentrations in winter were high, likely due to an overestimate in emissions from prescribed burning, fireplace combustion and yard waste burning throughout the winter. Tian et al. (Tian, *et al.*, 2008) found similar levels of errors in fire emissions.

Zinc (Zn) comes mainly from mobile and industrial sources. Simulated levels are low at most of the sites in CMAQ/LGO results (i.e., the smallest NMFB is -28% at JST in July 2001, and the largest is -160% at BHM in both periods). However, the simulated levels are overestimated in CMAQ/RG results (a NMFB of 51% at JST and 150% at

BHM). The performance of Zn simulation differs between CMAQ/LGO and CMAQ/RG because RG has a source profile for industrial sources that is high in Zn (6% of primary PM2.5) while the LGO study does not account for this. On the other hand, EC, which is an important tracer for diesel engine exhaust, was underestimated at all sites in both CMAQ/RG and CMAQ/LGO (Figure 2.3) except for JST, suggesting that regional emissions from mobile sources are largely underestimated in most of the SEARCH sites except for JST.

Al and Si, which are derived primarily from soil dust, have wintertime observations that are about one third of summer levels. However, CMAQ results show larger winter values, being biased quite high with NMFBs of over 100%. This suggests that the seasonal variation in $PM_{2.5}$ crustal emissions is not captured well. Reducing estimated emissions from soil dust by 50% or more would bring simulated and observed values in line. In CMAQ simulations, unpaved road dust and construction dust are the two major sources of Al and Si. Road dust is likely to be less seasonally dependent than dust from construction activities. Money spent for construction in January is about 70% of that in July, but this temporal variation is not considered in a SMOKE temporal profile. If we consider this temporal variation, it decreases emissions by 20% in January and increases emissions by 10% in July. In addition, errors in source profiles could be another reason for disagreement. Marmur et al. (2007) decreased the ratio of Al to $PM_{2.5}$ significantly in their optimized source profiles for JST. They observed that the modified source profile better matched measurements in Atlanta. Mn and Fe also come, in part, from soil dust, but mostly from other sources, and the simulated Mn and Fe agree fairly well with the observations, especially in the summer (NMFBs of Fe and Mn are around 50%).

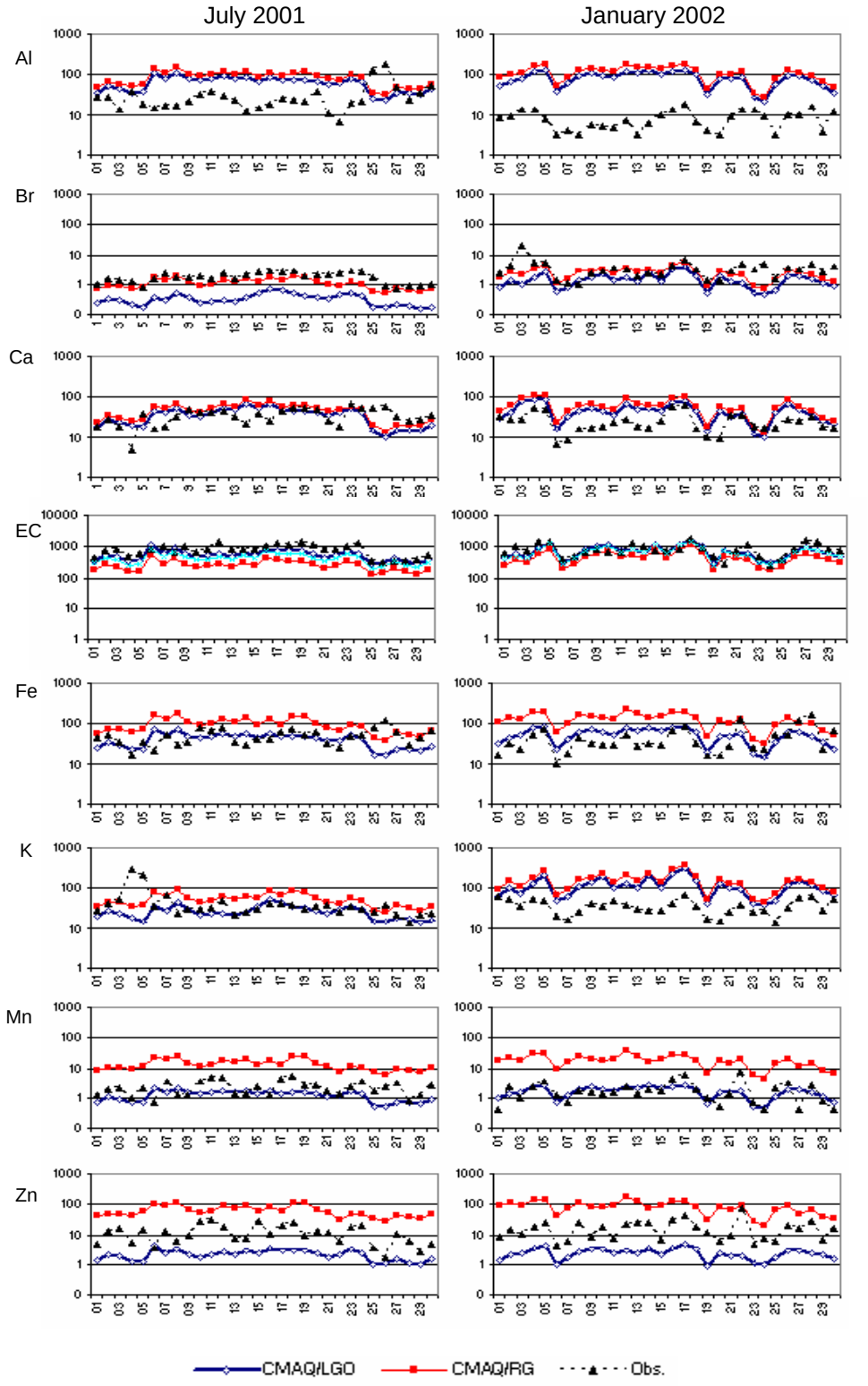


Figure 3.2 Comparison of measured and simulated metal concentrations. Values are daily averaged concentrations of SEARCH sites in $\mu g/m^3$. The y-axis is on a log scaled $\mu g/m^3$ unit, and the x-axis is day of month.

Table 3.3 Normalized performance of CMAQ simulations using source profiles in LGO (CMAQ/LGO) and in RG (CMAQ/RG)

metal	No. of records	mean OBS (ngm^{-3})	mean CMAQ (ngm^{-3})		MFB (%)		MFE (%)	
July 2001								
			CMAQ/ LGO	CMAQ/ RG	CMAQ/ LGO	CMAQ/ RG	CMAQ/ LGO	CMAQ/ RG
AL	195	39	61	122	50	61	98	108
As	227	0	0	45	191	199	191	199
Ba	227	19	1	2	-168	-156	168	156
Br	227	2	1	1	-119	-64	126	95
Ca	227	37	35	52	8	16	68	69
Fe	227	50	36	104	-18	19	66	74
K	227	54	26	52	-29	4	69	67
Mn	151	3	1	17	-63	49	77	98
Se	100	1	1	2	-39	-15	67	75
Si	227	83	168	308	69	77	105	111
Zn	227	12	2	55	-95	60	-94	102
January 2002								
AL	110	11	96	184	140	146	147	154
As	185	0	0	69	194	199	195	199
Ba	185	15	2	3	-163	-146	163	146
Br	185	4	2	3	-71	-37	91	82
Ca	185	26	48	72	61	72	87	94
Fe	185	42	52	150	40	76	77	100
K	185	36	111	150	64	80	91	105
Mn	91	4	2	30	-34	90	65	122
Pb	65	10	0	38	-168	25	168	123
Se	80	2	1	2	-33	4	58	73
Si	185	33	243	437	140	146	146	152
Zn	185	18	2	83	-130	41	131	102

3.3.3 Scaling factors for emissions

Different sets of scaling factors are calculated using the CMAQ/LGO (F_{LGO}) and the CMAQ/RG (F_{RG}) results, using ridge regression (equation (3)) to reduce differences between simulated and observed concentrations. F_{LGO} and F_{RG} agree on the direction of changes, but differ in quantity (Figure 3.4). Since RG and LGO use the same wood combustion, coal combustion and soil dust profiles (Chow *et al*., 2004), the differences in the mobile source profiles and the metal processing profiles lead to differences in the scaling factors. Both F_{LGO} and F_{RG} suggest that estimated emissions from soil dust are two to 30 times high (the scaling factors range from 0.03 in January to 0.5 in July). These huge seasonal changes in $F_{LGO,dust}$ and $F_{RG,dust}$ are likely not due to meteorological factors, since the NMFB of SO_2 and SO_4 increase slightly in January. Possible reasons are larger biases in emission estimates and removal mechanisms of dust in January versus July. Overestimation of fugitive dust emissions has been highlighted before (Pace, 2005). The lack of a nearby source removal mechanism, which is dependent on surface winds, as well as incorrect silt content data, may be responsible for the biases in CMAQ simulation (Roe and Hemmer, 2004).

Emissions from wood combustion appear low by 30% (JST and urban areas) to 60% (rural areas) in July. In January, emissions seem unbiased at urban sites, and appear high by 50 to 80% in other areas. Seasonal and location variations in the wood combustion scaling factors are possibly related to changes in dominant sources. In CMAQ-TR, wood combustion has six sub-categories: prescribed burning, wild fires, fireplaces and wood stoves, residential land clearing debris and yard waste burning (residential waste burning), agricultural burning, and wood/bark combustion in industrial

sectors. In July, residential waste burning is the major contributor at all sites except OLF, where wild fires are the largest contributor to the wood combustion category, and JST. Fireplaces and wood stove emissions are simulated to dominate biomass burning-related $PM_{2.5}$ at JST (0.1 $\mu g/m^3$ from total wood combustion), but contribute less than 5% of the total primary $PM_{2.5}$ (2.8 $\mu g/m^3$) (Figure 3.5). In winter, fireplace and woodstove emissions are the major wood burning sources over most areas except GFP where residential waste burning is the major contributor, and CTR where prescribed burning is dominant. At JST, fireplaces and woodstoves contribute to 80% of the simulated total wood combustion impact. Thus, the suggested reduction of emissions by 60% from the regression analysis shows that emissions from fireplaces and wood stoves are overestimated. Highly biased emission estimates may come from an excessive amount of regional wood consumption or over-allocated emissions to the urban area. In the EPA NEI, 30% of wood consumption from woodstoves is apportioned to urban areas and 70% to rural areas (E.H. Pechan & Associates Inc., 2006), but the spatial distribution of allocated emissions used in CMAQ is closer to the population density distribution than to residential wood usage for heating (Figure 3.3). The usage of fireplaces is likely to follow population density, but woodstoves less so. Allocating statewide woodstove emissions by residential fuel usage of wood, which is used in SMOKE to allocate county level wood stove and fireplaces emissions into grids may reduce biases (EPA, 2004).

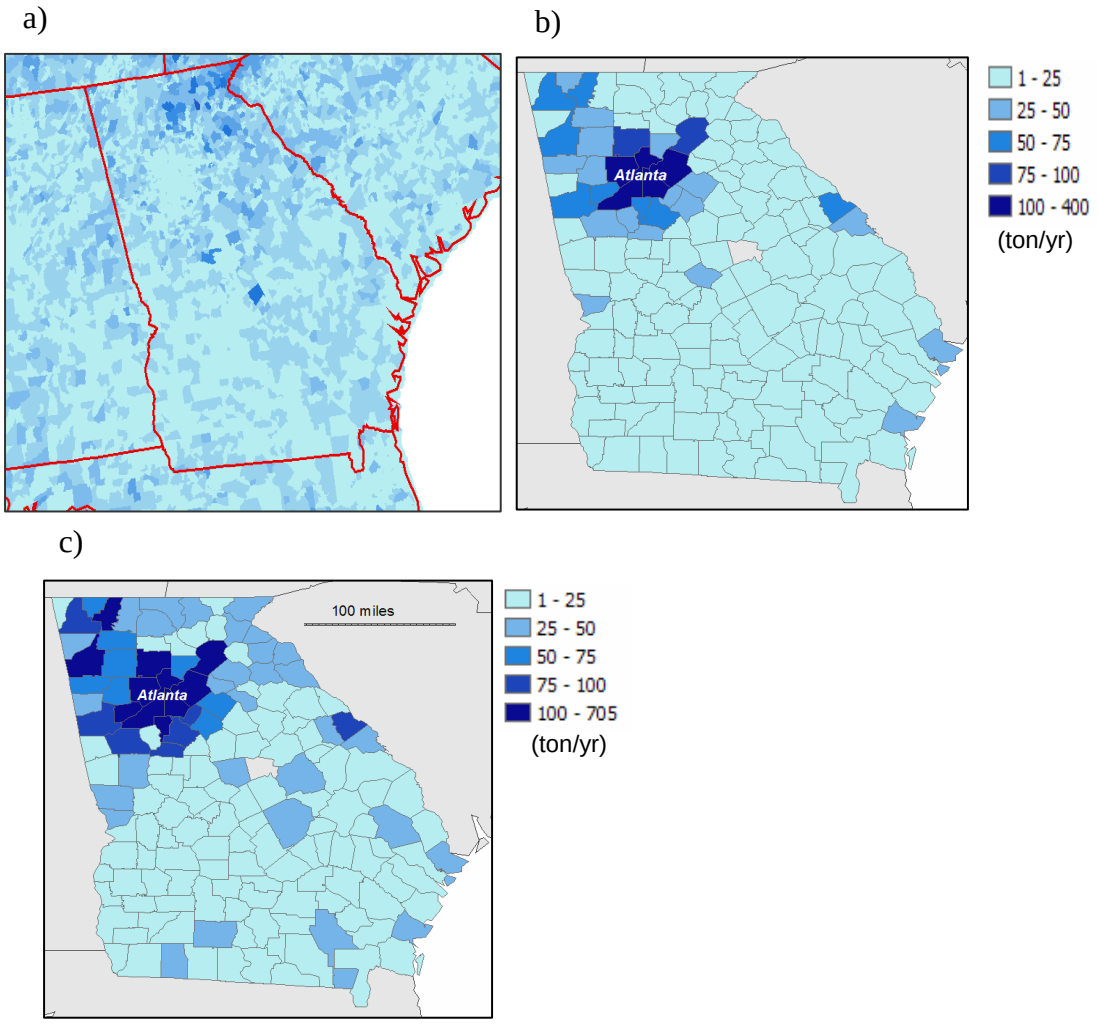


Figure 3.3 Spatial distribution of a) residential fuel usage – wood, b) $PM_{2.5}$ emissions from fireplaces in each county, c) $PM_{2.5}$ emissions from woodstoves in each county in 2001. Unit of emission is tons per year.

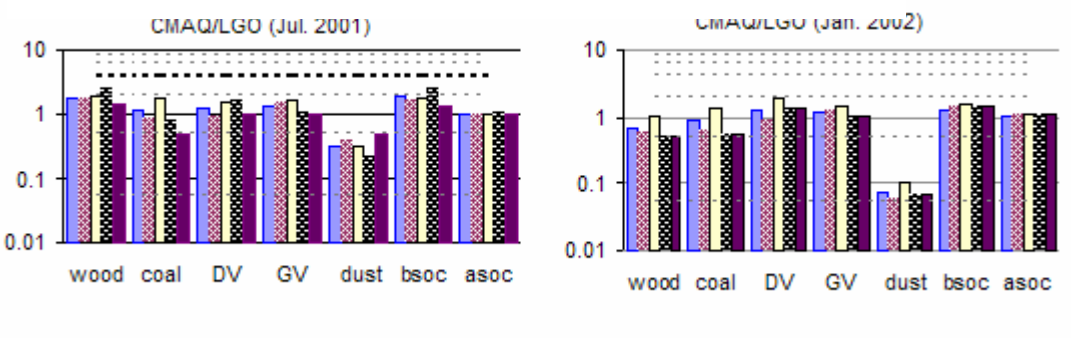
F_{LGO} and F_{RG} for mobile sources indicate emissions are likely biased low except for $F_{LGO,\ diesel}$ at JST. This is expected because simulated EC is significantly low at most of the SEARCH sites, and diesel engine exhaust is believed to be the major contributor to ambient EC levels. Underestimates in mobile source emissions come from two sources: 1) emission activity, i.e., vehicle miles traveled (VMT) of each vehicle type, and 2) emission factors. Emission factors are more likely biased in the case of gasoline engine exhaust, considering good CO performance in CMAQ. CO mainly comes from gasoline vehicles, and the NMFB for CO at SEARCH sites is slightly biased high (about 22% in July and 7% in January), which, in turn, suggests that gasoline vehicle VMT is at least not significantly underestimated. Emission estimates in MOBILE 6.0 are sensitive to many parameters, such as average speed, and the age of vehicles (Giannelli, *et al.*, 2002). Reasons for errors in emission factors may vary site by site and may need to be adjusted based on regions.

Zn is an important tracer for gasoline vehicles and metals processing (site-specifically, such as BHM). LGO study (Marmur et al., 2002) did not include a source profile for metals processing so does not CMAQ/LGO. Scaling factors for mobile sources in CMAQ/LGO results might be overestimated to compensate for the lack of industrial process emissions, especially at BMH where metal processes contribute largely (7% of $PM_{2.5}$ in both seasons) according to CMAQ source apportionment results. If a metal process source profile is added to LGO ($CMAQ/LGO_{metal}$), it decreases $F_{LGO,gasoline}$ about 20 to 25% in urban areas (Figure 3.4). This result highlights the problem of missing sources, especially point sources, in CMB modeling,. Point sources can have significant impacts depending on meteorology, though it is not easy to include all types of point

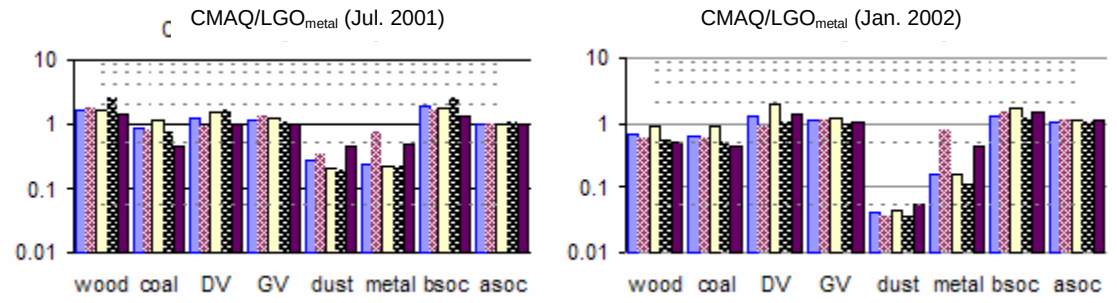
sources in CMB. Detailed source impacts simulated in CMAQ can provide information on possible important sources of $PM_{2.5}$ at a given monitoring site, help to select what source categories should be included in CMB studies, and identify potential biases in results of CMB models.

The scaling factors are for primary non-reactive pollutants, of which ambient concentrations are directly proportional to their emissions, and it is not clear that they should, or should not be, applied for secondary organic aerosol. However, possible incremental changes in biogenic secondary organic carbon (BSOC) and anthropogenic secondary organic carbon (ASOC) are studied here to evaluate potential improvement in BSOC and ASOC in CMAQ simulations. Simulated daily BSOC and ASOC concentrations are added as two additional sources with a source profile with only OC. Adding BSOC and ASOC sources has negligible impact on the scaling factors for the other source categories. Based on the regression analysis results, BSOC needs to be increased by 50% to 100% in July and 20% to 70% in January. Changes in ASOC are minimal since the F_{ASOC} are close to 1.0 in both seasons at all monitoring sites.

a) F_{LGO} using the original set of source profiles in Marmur et al. (2005)



B) A source profile of metal process is added to the LGO source profile



C) F_{RG}

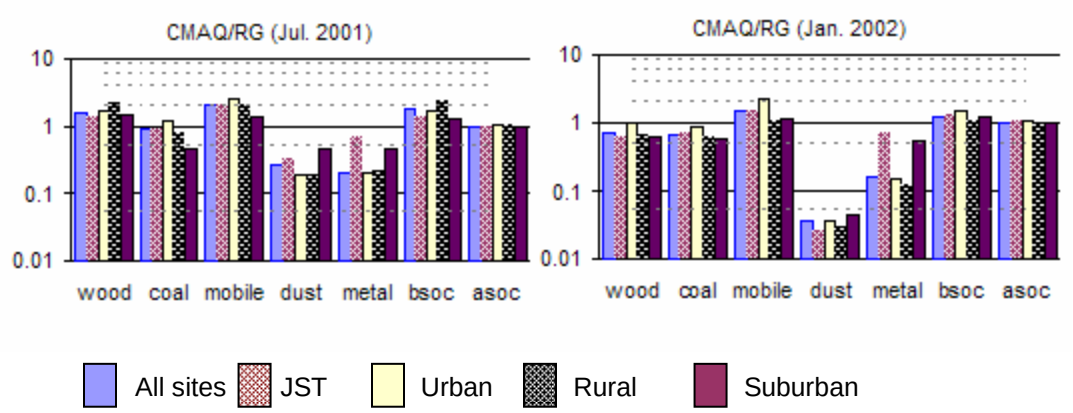


Figure 3.4 Scaling factors for source categories a) factors calculated using CMAQ/LGO results; b) factors calculated using the set of source profiles from the LGO study with a profile of metal processes, and c) factors using CMAQ/RG. Wood: biomass burning, coal: coal combustion, mobile: mobile sources, dust: soil dust, metal: metal processes, bsoc: biogenic secondary organic carbon, asoc: anthropogenic secondary organic carbon

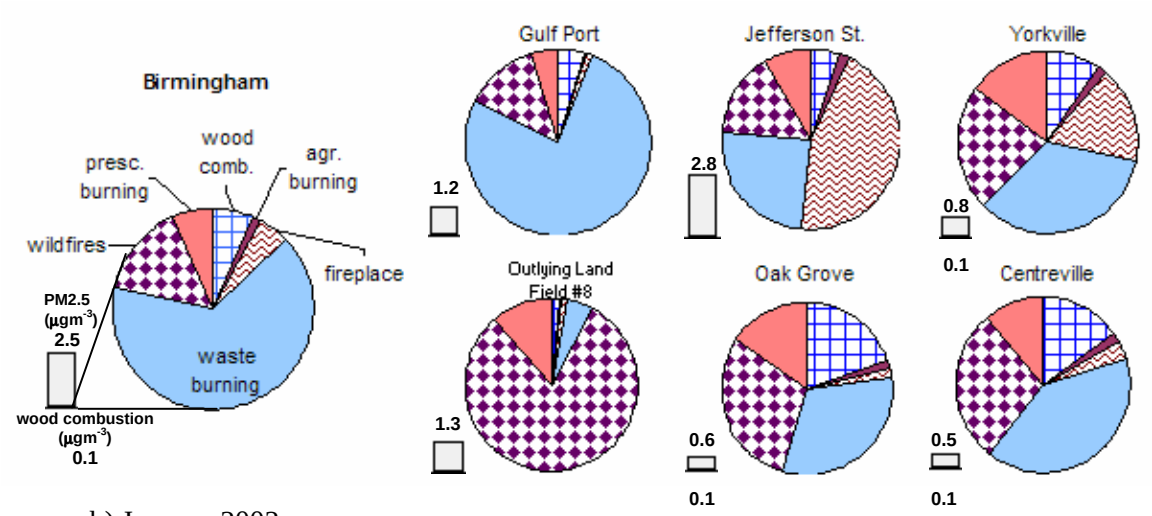
Table 3.4 Values and 95% confidence interval of scaling factors (F_{RG}, Figure 3.4)

	all sites	Jefferson St.	Urban area	Rural area	Suburban area
F_{RG}	(July)				
Wood burning	1.6 ± 0.22	1.4 ± 0.10	1.7 ± 0.20	2.2 ± 0.08	1.5 ± 0.48
Metal processes	0.2 ± 2.36	0.7 ± 0.66	0.2 ± 4.20	0.2 ± 1.02	0.5 ± 0.26
Mobile sources	2.1 ± 0.28	2.1 ± 0.60	2.6 ± 0.26	2.1 ± 0.10	1.4 ± 0.10
Coal combustion	0.9 ± 0.38	0.9 ± 0.44	1.2 ± 0.32	0.8 ± 0.40	0.5 ± 0.32
Soil dust	0.3 ± 1.64	0.3 ± 2.50	0.2 ± 1.76	0.2 ± 1.34	0.5 ± 0.96
Biogenic SOA	1.8 ± 0.30	1.4 ± 0.42	1.7 ± 0.34	2.5 ± 0.26	1.3 ± 0.14
Anthropogenic SOA	1.0 ± 0.02	1.0 ± 0.02	1.0 ± 0.02	1.1 ± 0.01	1.0 ± 0.01
F_{RG},	(January)				
Wood burning	0.8 ± 0.80	0.6 ± 1.26	1.0 ± 0.72	0.7 ± 0.70	0.6 ± 0.72
Metal processes	0.2 ± 2.60	0.7 ± 0.70	0.2 ± 4.42	0.1 ± 1.62	0.6 ± 0.36
Mobile sources	1.5 ± 0.18	1.6 ± 0.38	2.3 ± 0.16	1.1 ± 0.08	1.1 ± 0.06
Coal combustion	0.7 ± 0.42	0.7 ± 0.48	0.9 ± 0.36	0.6 ± 0.44	0.6 ± 0.40
Soil dust	0.0 ± 3.95	0.0 ± 6.12	0.0 ± 3.50	0.0 ± 3.62	0.0 ± 2.78
Biogenic SOA	1.2 ± 0.24	1.3 ± 0.22	1.5 ± 0.20	1.1 ± 0.26	1.2 ± 0.16
Anthropogenic SOA	1.1 ± 0.02	1.1 ± 0.06	1.1 ± 0.02	1.0 ± 0.04	1.0 ± 0.02

Table 3.5 Values and 95% confidence interval of scaling factors (F_{LGO} and $F_{LGO,metal}$ in Figure 3.4)

	all sites	Jefferson St.	Urban area	Rural area	Suburban area
F_{LGO}	(July)				
Wood burning	1.8 ± 0.24	1.9 ± 0.09	2.0 ± 0.24	2.6 ± 0.10	1.5 ± 0.51
Coal combustion	1.2 ± 0.44	0.9 ± 0.45	1.8 ± 0.49	0.8 ± 0.40	0.5 ± 0.31
Diesel exhaust	1.3 ± 0.56	1.0 ± 1.11	1.5 ± 0.57	1.7 ± 0.23	1.0 ± 0.25
Gasoline exhaust	1.3 ± 0.08	1.6 ± 0.18	1.6 ± 0.06	1.1 ± 0.01	1.1 ± 0.02
Soil dust	0.3 ± 1.80	0.4 ± 2.51	0.3 ± 2.15	0.2 ± 1.40	0.5 ± 1.00
Biogenic SOA	1.9 ± 0.32	1.7 ± 0.44	1.7 ± 0.36	2.6 ± 0.27	1.4 ± 0.15
Anthropogenic SOA	1.0 ± 0.02	1.1 ± 0.01	1.0 ± 0.01	1.1 ± 0.01	1.0 ± 0.01
F_{LGO}	(January)				
Wood burning	0.7 ± 0.85	0.6 ± 1.34	1.0 ± 0.80	0.6 ± 0.71	0.5 ± 0.75
Coal combustion	0.9 ± 0.53	0.7 ± 0.52	1.4 ± 0.62	0.6 ± 0.51	0.6 ± 0.43
Diesel exhaust	1.3 ± 0.36	0.9 ± 0.75	1.9 ± 0.35	1.1 ± 0.17	1.3 ± 0.12
Gasoline exhaust	1.2 ± 0.06	1.3 ± 0.13	1.5 ± 0.04	1.0 ± 0.02	1.1 ± 0.02
Soil dust	0.1 ± 4.41	0.1 ± 6.05	0.1 ± 4.75	0.0 ± 2.91	0.1 ± 2.93
Biogenic SOA	1.3 ± 0.24	1.5 ± 0.24	1.6 ± 0.22	1.2 ± 0.28	1.5 ± 0.18
Anthropogenic SOA	1.1 ± 0.03	1.1 ± 0.05	1.1 ± 0.01	1.0 ± 0.03	1.1 ± 0.02
$F_{LGO,metal}$	(July)				
Wood burning	1.6 ± 0.24	1.5 ± 0.10	1.7 ± 0.22	2.4 ± 0.09	1.4 ± 0.52
Coal combustion	0.9 ± 0.38	0.8 ± 0.45	1.2 ± 0.33	0.7 ± 0.39	0.4 ± 0.32
Diesel exhaust	1.2 ± 0.56	0.9 ± 1.11	1.5 ± 0.57	1.6 ± 0.24	1.0 ± 0.26
Gasoline exhaust	1.1 ± 0.08	1.2 ± 0.18	1.2 ± 0.06	1.1 ± 0.01	1.0 ± 0.02
Soil dust	0.2 ± 1.68	0.3 ± 2.49	0.2 ± 1.84	0.2 ± 1.39	0.5 ± 1.00
Metal processes	0.2 ± 1.40	0.4 ± 0.35	0.2 ± 2.55	0.2 ± 0.48	0.5 ± 0.14
Biogenic SOA	1.8 ± 0.32	1.4 ± 0.43	1.8 ± 0.36	2.5 ± 0.28	1.4 ± 0.16
Anthropogenic SOA	1.0 ± 0.02	1.0 ± 0.02	1.0 ± 0.02	1.1 ± 0.01	1.0 ± 0.00
$F_{LGO,metal}$	(January)				
Wood burning	0.5 ± 0.85	0.5 ± 1.33	0.6 ± 0.38	0.5 ± 0.37	0.4 ± 0.75
Coal combustion	0.5 ± 0.45	0.4 ± 0.51	0.7 ± 0.19	0.4 ± 0.24	0.4 ± 0.42
Diesel exhaust	1.2 ± 0.36	0.9 ± 0.73	1.9 ± 0.18	1.0 ± 0.09	1.4 ± 0.12
Gasoline exhaust	1.0 ± 0.06	1.1 ± 0.13	1.1 ± 0.02	1.0 ± 0.01	1.1 ± 0.01
Soil dust	0.0 ± 4.08	0.0 ± 6.03	0.0 ± 1.85	0.0 ± 1.92	0.1 ± 2.93
Metal processes	0.2 ± 1.83	0.2 ± 0.46	0.2 ± 1.57	0.1 ± 0.58	0.4 ± 0.28
Biogenic SOA	1.3 ± 0.26	1.4 ± 0.24	1.5 ± 0.11	1.2 ± 0.14	1.5 ± 0.18
Anthropogenic SOA	1.1 ± 0.04	1.1 ± 0.06	1.1 ± 0.01	1.0 ± 0.02	1.1 ± 0.02

a) July 2001



b) January 2002

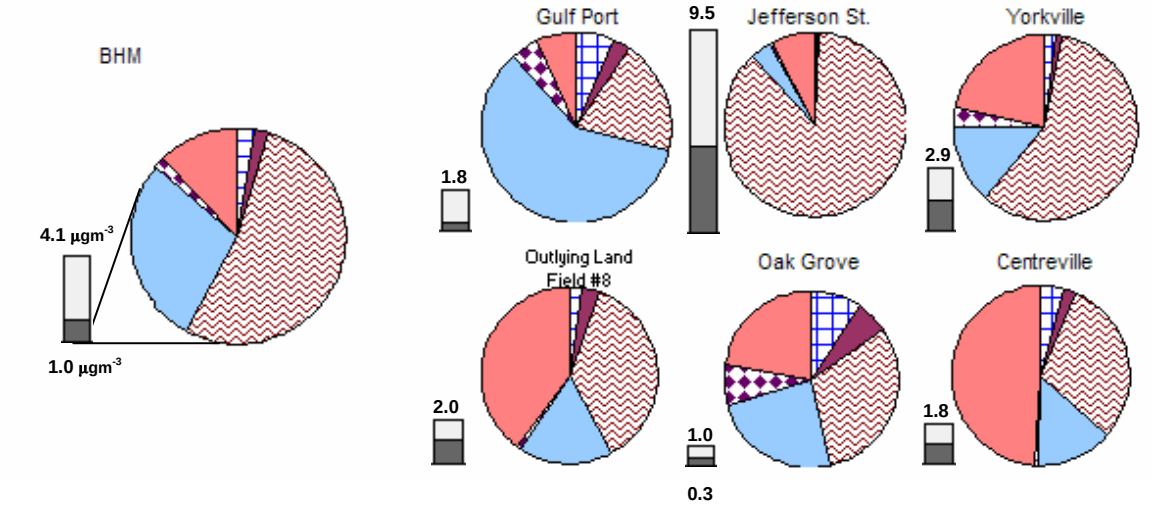


Figure 3.5 Contributions of wild fire, prescribed burning (presc. burning), wood/bark combustion (wood comb.), agricultural burning (agr. burning), fireplaces and woodstoves (fireplace), and residential yard waste burning/land clearing debris (waste burning) to total wood combustion. The bar represents total primary $PM_{2.5}$ concentrations from all wood combustion (dark grey) and the rest of sources (white) Numbers on the top of bars are total primary $PM_{2.5}$ concentrations and numbers at the bottom are total biomass burning-related $PM_{2.5}$ concentrations in $\mu g/m^3$.

3.3.4 Performance of adjusted $PM_{2.5}$ levels

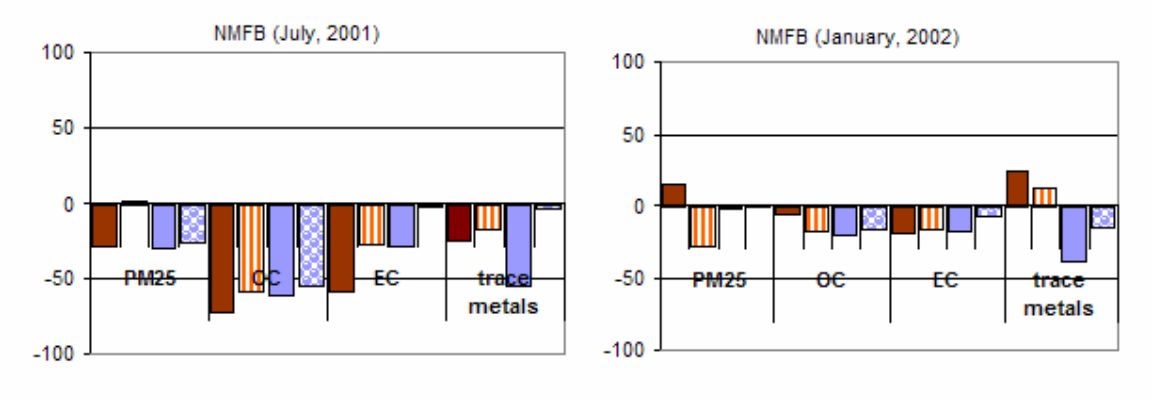
To evaluate how much the performance of CMAQ simulations would be improved by applying the scaling factors, simulated $PM_{2.5}$ concentrations are multiplied with corresponding scaling factors. The original simulated $PM_{2.5}$ at a site k equals the sum of $PM_{2.5}$ from all the source categories ($PM_{2.5jk}$). The adjusted $PM_{2.5,k}$ is the total of the multiplication of $PM_{2.5,jk}$ and f_{jk}. NMFBs and NMFEs of the adjusted $PM_{2.5}$ are calculated and compared with the original $PM_{2.5}$ (Figure 3.6). BSOC and ASOC in CMAQ are not adjusted with the scaling factors, since SOC is not directly proportional to emissions. Overall, applying the scaling factors typically improved CMAQ performance. However, the NMFE of $PM_{2.5}$ in July increases by 10% after applying F_{LGO}, and the NMFB of OC worsens when $F_{LGO,metal}$ and F_{RG} are applied. Scaling factors are calculated using the weighted summation of errors for OC, EC and other trace metals, but this does not guarantee improvement in NMFE or NMFB of individual components of $PM_{2.5}$. The improvement in trace metal simulations is not as good as for $PM_{2.5}$ because the scaling factors are estimated to minimize the weighted summation of errors.

Using F_{LGO} rather than F_{RG} results in better improvement in $PM_{2.5}$ simulations in January, and vice versa in July. Site-wide OC performance improves from -80% to -50% (July 2001), and from -20% to -8% (January 2002). Improvement in OC simulations is better in January than in July, largely because there is less secondary organic aerosol production in wintertime. Previous studies suggested that more than approximately 50% of OC in summer in the southeastern US comes from biogenic sources (Lemire, *et al.*, 2002;Lewis, *et al.*, 2004), and underestimates of OC in air quality models are due, in part, to underestimates of biogenic SOA formation. Regression analysis suggests that BSOC in

CMAQ should be increased more than 100% in both winter and summer. EC has only primary sources, and EC performance improved significantly in both seasons, but simulations are still lower than observations.

Metal profiles that are used here do not cover all the $PM_{2.5}$ sources. For example, the largest contributor missing a source profile is meat cooking (contributes 5 to 10% of $PM_{2.5}$). Metal emissions from meat cooking are minor (the ratio of total trace metals to $PM_{2.5}$ that emitted from meat cooking is less than 2%) (Hildemann, *et al.*, 1991), so this issue primarily contributes to EC and OC biases. Missing industrial processes (other than metal processes) contribute 6% (July) to 10% (January) of $PM_{2.5}$ and may have large impact on metal simulations, Based on the source profiles listed in SPECIATE 4.0, industrial processes are important sources of calcium, potassium, copper, iron and silicon (EPA, 2006).

a) Normalized mean fractional bias



b) Normalized mean fractional error

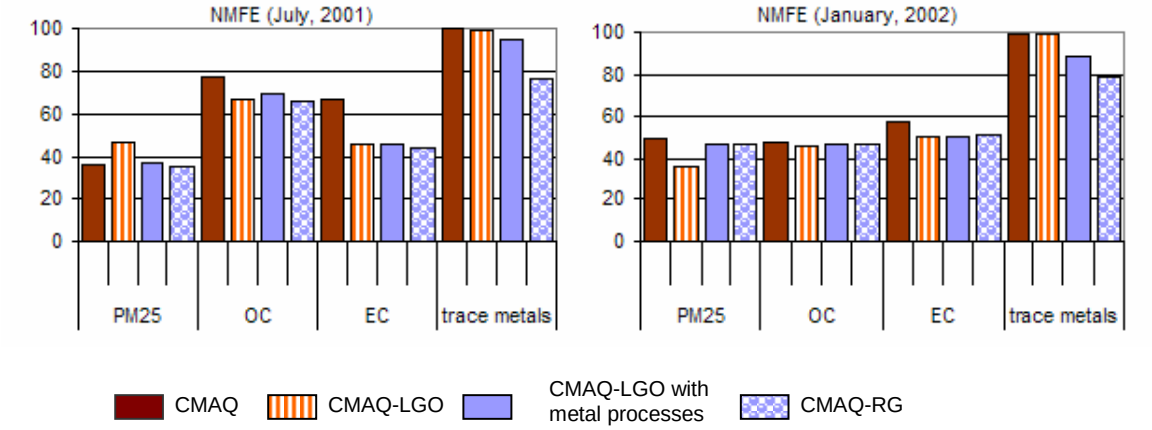


Figure 3.6 Changes in normalized mean fractional biases (a) and errors (b) for $PM_{2.5,}$ OC, EC, and trace metals at SEARCH monitoring sites. CMAQ; the original CMAQ performance. LGO; the changed performance after applying F_{LGO}. RG; the changed performance after applying $F_{RG.}$

3.3.5 Source impacts of trace metals

Applying scaling factors to $PM_{2.5}$ source impacts also improves trace metal estimation (Figure 3.7). Coal fired power plants, wood combustion and metal processing are the major sources of trace metals. Gasoline engine exhaust and diesel engine exhaust are important to EC. Adjusted Al and Si estimates matched observations better than the CMAQ base case, but possibly over-correct in January since adjusted Al and Si simulations are less than half of observations.

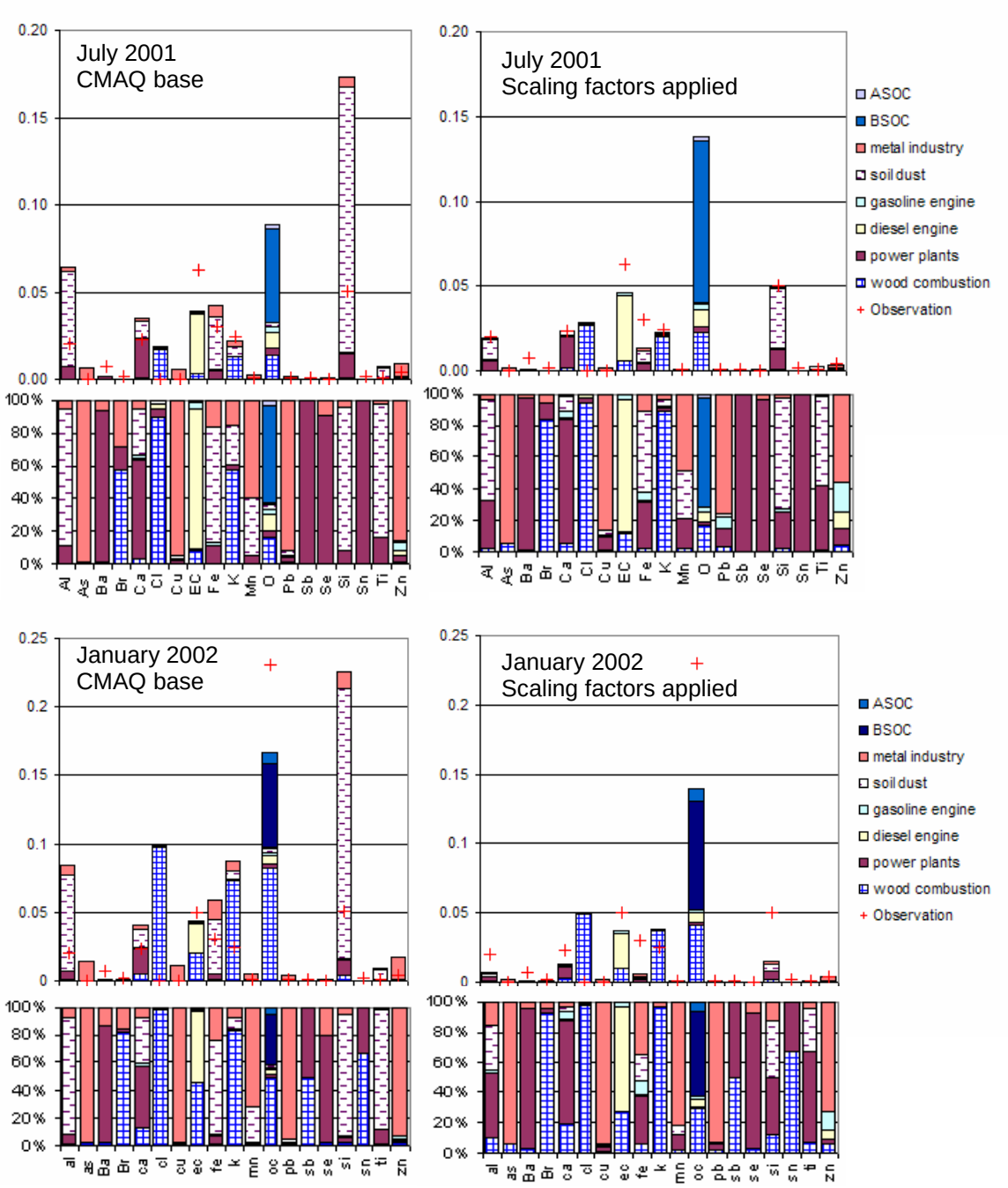


Figure 3.7 Source apportionment of trace metals in July 2001 and January 2002. Note that EC and OC bars are reduced to one tenth of original concentrations for display. Scaling factors are not applied to secondary organic carbon (ASOC: anthropogenic SOC, BSOC: biogenic SOC). SOCs are included in plots to compare with observations.

3.3.6 Sensitivity analysis of scaling factors

Sensitivity analysis of scaling factors ($F_{all,LGO}$) to a specific species is tested by brute force method. Scaling factors for wood burning have different sensitivities in different seasons. In July, Br, EC, and OC are major species that determine the scaling factor while K dominates in winter. Si primarily determines the scaling factor for soil dust, and Zn for metal processes. EC and OC are important species estimating the scaling factors for all source categories.

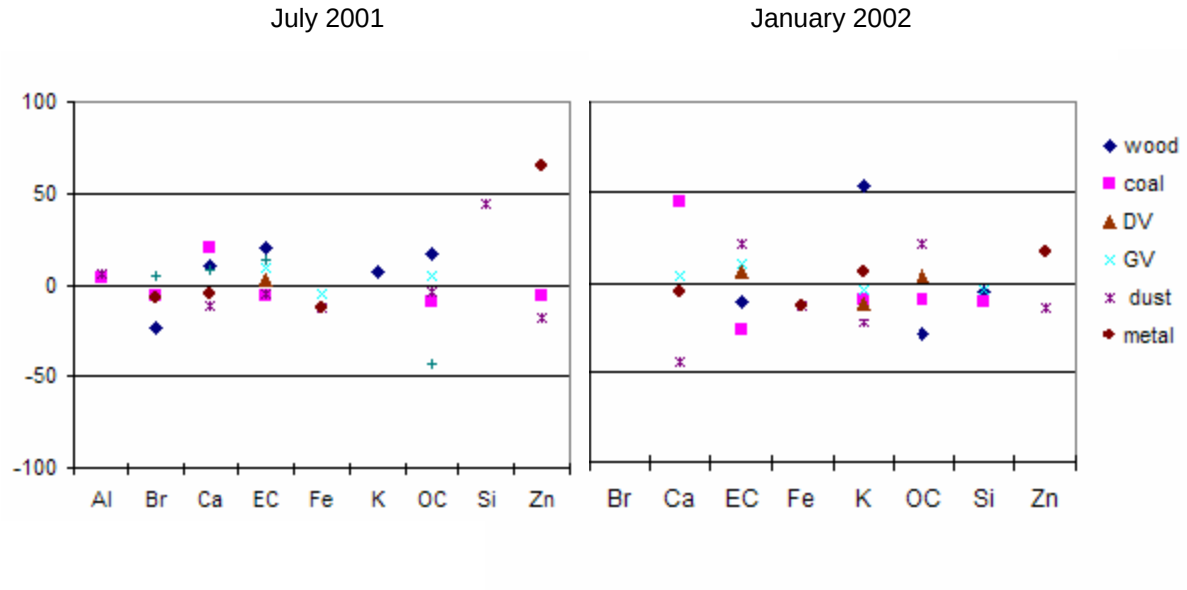


Figure 3.8 Normalized changes in scaling factors when a specific species is excluded from the regression analysis. For example, 10% of wood burning at X-axis = K indicates that a scaling factor of wood burning decreases 10% when potassium (K) is eliminated from data set. wood: wood burning, coal: coal combustion , DV: diesel vehicle, GV: gasoline vehicle, dust: fugitive dust, metal: industrial metal processes

3.4 CONCLUSION

Source apportionment of primary $PM_{2.5}$ using CMAQ with trace metals shows that while many metal species are reasonably well simulated, those species linked to biomass burning and mobile sources have seasonally- and spatially-dependent biases. Emission estimates from those species are suspected to be the reason for these biases based on the good performance found for simulated CO, SO_2 and ozone. Ridge regression analysis can quantify how biases in inventoried emissions lead to discrepancies between daily metal measurements at SEARCH sites and simulated levels. Results from our work show likely problems in emissions estimates and their spatial allocation, as well as potential errors in source profiles. Biases in emissions appear reduced when scaling factors are applied to CMAQ results. Emissions from woodstoves and fireplaces appear to be over-allocated to urban areas in the winter, and the quantity of emissions from residential yard waste burning is underestimated especially in rural areas in the summer. $PM_{2.5}$ emission factors for gasoline vehicles, VMT, and/or emission factors for diesel engine exhaust are likely to be underestimated. The soil dust contribution is biased high, possibly due to lack of a source removal mechanism in CMAQ, or not accounting for meteorological factors such as rain and wind. Missing source profiles for industrial processes or uncertainties in the source profiles leads to an increased discrepancy between simulations and measurements, especially for Zn. These results also suggest that having an incomplete specification of sources when applying CMB can lead to potential biases in source apportionment results, which is significant given that the typical application of CMB only includes sources of about 80% of the primary $PM_{2.5}$ emissions.

CHAPTER 4

ASSESSMING EMISSION INVENTORIES USING ORGANIC MOLECULAR MARKERS

4.1 INTRODUCTION

Receptor- and source-oriented air quality models (AQMs) are used to quantify air quality impacts from specific sources of fine particulate matter with a diameter smaller than $_{2.5}$ μm ($PM_{2.5}$). Receptor-oriented models (ROMs) infer source contributions by determining best-fit combinations of emission source profiles to reconstruct the measured chemical composition in the ambient air (Watson et al., 1984) at specific monitoring stations. Chemical mass balance (CMB) and positive matrix factorization (PMF) (Liu et al., 1982) are commonly used ROMs. CMB applications largely rely on metal tracers (CMB regular model, RG), though recent applications have also used organic molecular markers (Schauer, 2003; Zheng et al., 2002) and gaseous species (CMB with Lipschitz Global Optimizer, LGO) (Marmur et al., 2005). On the other hand, emission-based models, such as the Community Multi-Scale Air Quality (CMAQ) model (Byun and Ching, 1999), the comprehensive air-quality model with extensions (CAMx) (Yarwood et al., 2004), and the UCD/CIT air quality model (Held et al., 2005) rely on emission estimates to simulate domain-wide atmospheric concentrations.

Previous studies have shown that emission-based AQMs and ROMs give qualitatively consistent results, although significant differences in absolute impacts are found (Held et al., 2005; Lane et al., 2007; Lee et al., 2008; Marmur et al., 2006; Ying et

al., 2004). Different types of ROMs also find quantitatively dissimilar source impacts (Hopke et al., 2006). These discrepancies are partly caused by uncertainties inherent in the models and inputs. CMB modeling is limited, in part, by the accuracy of measured concentrations, source profiles, and completeness of sources treated (Cheng et al., 1988). Accuracy of emission-based models is limited by errors in emission inventories, meteorological data, and their ability to simulate physical and chemical processes in the atmosphere. These uncertainties can be addressed by combining AQMs and ROMs together (Chow and Watson, 2002).

In this chapter, simulated concentrations of organic molecular markers are compared with observations to identify biases in emission estimates. This work extends the previous work in Chapter 4 wherein trace metals are used for assessing $PM_{2.5}$ emission inventories. Since there are more tracer species available for organic molecular marker measurements compared to metal measurements (typically 60 species vs. 14 species), organic molecular markers can provide additional information on specific sources (e.g., wood burning and mobile sources) and distinguish additional source categories (e.g., meat cooking and natural gas combustion).

4.2 METHODS

4.2.1 CMAQ modeling and measurement data

Following the same method as described in detail in Chapter 2 and Chapter 3, and summarized below, source apportionment of $PM_{2.5}$ is conducted by applying the CMAQ tracer method (CMAQ-TR) to the continental United States and portions of Canada and Mexico for July 2001 and January 2002. Monitoring data from the Air Quality System

(AQS), (EPA, 2007) Speciated Trends Network (STN), Interagency Monitoring of Protected Visual Environments (IMPROVE, 1995), Assessment of Spatial Aerosol Composition in Atlanta (ASACA) (Butler et al., 2003), and SEARCH are used. The SEARCH (EPRI, 2007) network includes daily measurements at Jefferson St. (JST) in Atlanta, Georgia, and monthly-averaged measurements at the other sites (i.e., Birmingham (BHM) and Centreville (CTR) in Alabama, Gulfport (GFP) and Oak Grove (OAK) in Mississippi, Outlying Land Field #8 (OLF) in Florida, and Yorkville (YRK) in Georgia) (Zheng et al., 2006). Most of the analysis here is focused on JST, which has more detailed observation data than other SEARCH sites. Emission estimates used in CMAQ come from the EPA national emission inventory (NEI) 2001, updated with emissions from biomass burning in the Visibility Improvement State and Tribal Association of the Southeast (VISTAS) in 2002, and additional emission sources such as meat cooking, and cigarette smoking. Source apportionment results from CMAQ (28 source categories) are combined with source profiles used in molecular marker-based CMB (MM) in order to estimate concentrations of tracer species:

$$C_i = \sum_{j=1}^{n} S_{PM\,2.5,\,j} \times (R_{i,\,j} / R_{PM\,25,\,j}) \qquad (4.1)$$

where $S_{PM2.5,j}$ is the impact of source j on $PM_{2.5,}$ $R_{i,j}$ is the emission rate of the organic molecular marker i from the source j, and $R_{PM2.5,j}$ is the emission rate of $PM_{2.5}$ from the source j. Species used include elemental and organic carbon, inorganic ions, aluminum (Al), silicon (Si) and organic molecular markers.

In MM modeling, emission rates of tracer species for seven primary emission sources are defined; meat cooking, wood burning, diesel engine exhaust, gasoline engine exhaust, natural gas combustion, road dust, and cigarette smoking. Cigarette smoking is

omitted from this study since there was no source contribution from cigarette smoking for the studied period in MM results. 28 source categories in CMAQ are grouped together to match the corresponding sources in CMB studies (Table 4.1). Some source categories such as fuel combustions, industrial processes and fugitive dust that are used in CMAQ do not have matching CMB source categories, thus it is excluded from regression analysis.

Table 4.1 Source categories in MM and its matching categories in CMAQ

MM	CMAQ	MM	CMAQ
Meat cooking	Meat cooking	Diesel engine exhaust	On-road diesel vehicles Non-road diesel engines
Natural gas combustion	Natural gas combustion – boilers Natural gas combustion – residential	Gasoline engine exhaust	On-road gasoline vehicles Non-road gasoline engines
Wood burning	Wild fire Prescribed burning Fireplaces/woodstoves Residential yard waste burning	Road dust	Paved road dust Unpaved road dust
Vegetative detritus	Not available		

Table 4.2 Organic molecular markers and its major sources used in MM

Source Category	Organic molecular marker species
Wood burning	Levoglucosan, resin acids, retene, n-alkanes, PAHs
Vehicle exhaust	Hopanes, steranes, PAHs, n-alkanes
Natural gas combustion	n-alkanes
Vegetative detritus	n-alkanes
Road dust	Aluminum, silicon

Vegetative detritus used in MM is not included in CMAQ modeling because emission data are not available. Mobile source profiles (i.e., diesel engine exhaust and gasoline engine exhaust) by Rogge et al. (1993) were recently updated by Lough et al. (2007) (SP_{Lough}). Two analyses are conducted using both versions of the mobile source profiles: $SP_{Rogge,}$ which is the set of seven source profiles with mobile source profiles from Rogge et al.'s study (1993); and SP_{Lough} , which is the similar set of source profiles, but includes Lough et al's update (2007).

4.2.2 Regression analysis

Simulated organic molecular markers differ from measured values, and the emission inventories are assumed to be the major reason for these differences since model performance for gaseous species suggests that errors in meteorological and photochemical processes are relatively minor contributors to PM concentration biases. To estimate potential biases in emission estimates, scaling factors for each source category are calculated using a ridge regression method with a length parameter to minimize changes in scaling factors. In matrix form, the scaling factors, **F**, are calculated as

$$\mathbf{F} = (\mathbf{X}^T\mathbf{W}\mathbf{X} + \lambda\mathbf{E})^{-1}(\mathbf{X}^T\mathbf{W}\mathbf{Y} + \lambda\,\mathbf{[1]}) \qquad (4.2)$$

where **F** is an S by 1 vector (S is the number of sources and f_j is an element of F and a scaling factor for source j). **X** is an N by S matrix (where N equals the number of species,J times number of measurement periods) of simulated concentrations in CMAQ. **Y** is an N by 1 vector of observation data, **E** is an S by S identity matrix, [1] is an S by 1 column vector of one, and λ is a length parameter that is set as the average of the two lowest diagonal values of the $\mathbf{X}^T\mathbf{W}\mathbf{X}$ term. Confident level of **F** is estimated using the

variance-covariance matrix **V** and the t-test value for 95% interval with number of records more than 150 (infinity):

$$\mathbf{V} = (\mathbf{X}^T\mathbf{W}\mathbf{X} + \lambda\mathbf{E})^{-1}(\mathbf{X}^T\mathbf{W}\mathbf{X})(\mathbf{X}^T\mathbf{W}\mathbf{X} + \lambda\mathbf{E}) \quad (4.3)$$

$$t_{2.5\%,infinity} = 1.96 \quad (4.4)$$

W is the weighting matrix that accounts the uncertainties related with Equation (2) as well as the differences in order of magnitudes of observations between species. For example, EC, Al and hopanes measurements are on the order of 1 μgm^{-3}, 10 ngm^{-3}, and 1 $ngm^{-3,}$ respectively. We need to weight EC and Al with the inverse of the order of their measurements, i.e., the variances of measured species, are added to evenly weigh each species in the regression analysis. Thus, the weighting factor, W, in equation (4.2) is defined as:

$$w_{ik} = \frac{1}{\mathrm{var}(y_i) + \mathrm{var}(e_{CMAQ}) + \sum_j \left(\mathrm{var}(a_{ij}) \times SA_{jk}^2\right)} \quad (4.5)$$

where var(y_i) is the variance of observed species i at all monitoring sites, var(e_{cmaq}) is the variance of CMAQ simulation error (i.e., variance of the difference between CMAQ simulation and observation). Var(a_{ij}) is the uncertainty of emission rates of species i in the source profile for the source category j. The simulated source impacts of a source j at a site k (SA_{jk}) vary daily and w_{ik} changes accordingly. e_{cmaq} for $PM_{2.5}$ is estimated in Chapter 3 and assumed to be the same as those of organic molecular marker simulations. The variance of errors in the $PM_{2.5}$ simulation is converted to those of molecular markers ($e_{cmaq,i}$) (Table 4.3) by multiplying the ratio of average molecular marker measurements to average $PM_{2.5}$ measurements:

$$\mathrm{var}(e_{CMAQ,i}) = \left(\frac{\bar{c}_i^{OBS}}{\bar{c}_{PM\,2.5}^{OBS}} \right)^2 \times \mathrm{var}(e_{CMAQ,PM\,2.5}) \qquad (4.6)$$

where the average of an organic molecular marker i ($\bar{c}_i^{OBS}$) $PM_{2.5}$ ($\bar{c}_{PM\,2.5}^{OBS}$) are in ngm^{-3}.

Table 4.3 Uncertainties in estimated organic molecular marker concentrations (units: $[\mu gm^{-3}]^2$)

July 2001 $var(e_{cmaq,PM2.5}) = 1.86\ (\mu gm^{-3})^2$				January 2002 $var(e_{cmaq,PM2.5}) = 9.8\ (\mu gm^{-3})^2$			
species	$var(e_i)$	species	$var(e_i)$	species	$var(e_i)$	species	$var(e_i)$
17a(H)-21b(H)-29-norhopane	0.003	tetracosanoic acid	0.142	17a(H)-21b(H)-29-norhopane	0.029	tetracosanoic acid	1.246
7h-benz(de)anthracen-7-one	0.007	pentacosanoic acid	0.002	7h-benz(de)anthracen-7-one	0.059	pentacosanoic acid	0.023
20R, aaa-cholestane	0.001	hexacosanoic acid	0.040	20R, aaa-cholestane	0.009	hexacosanoic acid	0.349
20R+S, abb-cholestane	0.000	heptacosanoic acid	0.003	20R+S, abb-cholestane	0.003	heptacosanoic acid	0.025
abietic acid	0.014	octacosanoic acid	0.081	abietic acid	0.127	octacosanoic acid	0.713
anteiso-triacontane	0.012	triacontanoic acid	0.097	anteiso-triacontane	0.108	triacontanoic acid	0.845
aluminum	255.75	17a(H)-21b(H)-hopane	0.006	aluminum	2250.44	17a(H)-21b(H)-hopane	0.051
benzo(b)fluoranthene	0.022	iso nonacosane	0.028	benzo(b)fluoranthene	0.194	iso nonacosane	0.247
benzo(e)pyrene	0.022	iso-hentriacontane	0.035	benzo(e)pyrene	0.186	iso-hentriacontane	0.306
benzo(ghi)perylene	0.038	indeno(cd)fluoranthene	0.011	benzo(ghi)perylene	0.338	indeno(cd)fluoranthene	0.096
benzo(k)fluroanthene	0.011	indeno(cd)pyrene	0.043	benzo(k)fluroanthene	0.096	indeno(cd)pyrene	0.383
nonanal	1.752	levoglucosan	4.487	nonanal	15.449	levoglucosan	39.503
cholesterol	0.004	pentacosane	0.208	cholesterol	0.039	pentacosane	1.825
20R+S, abb-ergostane	0.001	hexacosane	0.080	20R+S, abb-ergostane	0.004	hexacosane	0.697
hexadecenoic acid	0.188	heptacosane	0.015	hexadecenoic acid	1.652	heptacosane	0.134

Table 4.3 (continued)

July 2001 $var(e_{cmaq,PM2.5}) = 1.86\ (\mu gm^{-3})^2$				January 2002 $var(e_{cmaq,PM2.5}) = 9.8\ (\mu gm^{-3})^2$			
species	$var(e_i)$	species	$var(e_i)$	species	$var(e_i)$	species	$var(e_i)$
octadecenoic acid	0.279	octacosane	0.055	octadecenoic acid	2.454	octacosane	0.482
tetradecanoic acid	0.127	nonacosane	0.359	tetradecanoic acid	1.124	nonacosane	3.177
pentadecanoic acid	0.034	triacontane	0.027	pentadecanoic acid	0.296	triacontane	0.238
hexadecanoic acid	0.014	hentriacontane	0.461	hexadecanoic acid	0.120	hentriacontane	4.068
heptadecanoic acid	0.031	odotriacontane	0.034	heptadecanoic acid	0.276	odotriacontane	0.296
octadecanoic acid	2.511	tritriacontane	0.062	octadecanoic acid	22.088	tritriacontane	0.549
nonadecanoic acid	0.003	tetracontane	0.022	nonadecanoic acid	0.029	tetracontane	0.194
eicosanoic acid	0.049	propionylsyringol	0.000	eicosanoic acid	0.431	propionylsyringol	0.003
heneicosanoic acid	0.002	silicon	1353.9	heneicosanoic acid	0.023	silicon	11914.9
docosanoic acid	0.051	20R+S, abb-sitostane	0.002	docosanoic acid	0.444	20R+S, abb-sitostane	0.017
		17a(H)-21b(H)-hopane	0.000	tricosanoic acid	0.079	17a(H)-21b(H)-hopane	0.003

Sensitivity analysis of scaling factors to each organic molecular markers has been done to test the suitability of using the specific species as tracers and their importance to scaling emission inventories. The organic molecular markers are chosen from non-reactive species in the atmosphere. However, some studies have shown that they may be reactive, which would introduce biases in regression analysis results. It is also of interest to evaluate the importance of each molecular marker species to understand discrepancies between simulations and observations or scaling factors. Sensitivity of scaling factors is estimated based on changes in scaling factors when one species is excluded. For example, when elemental carbon (EC) is excluded from the calculation, **X** in the equation (4.2) becomes the ((J-1) × D) by S matrix, **Y**, the ((J-1) × D) vector. Normalized changes in scaling factors before and after omitting EC are calculated in order to estimate the impact of EC to the scaling factors.

4.3. RESULTS AND DISCUSSION

Model performance has been evaluated in Chapter 2 using monitoring data from AQS, STN, IMPROVE, ASACA, and SEARCH. Briefly, model performance for gaseous species (i.e., CO, SO_2 and O_3) was good, suggesting that meteorology data and photochemistry processes are well captured. Simulated inorganic aerosol is in acceptable ranges using metrics from Boylan et al. (2003), but organic carbon is usually underestimated at most locations, and has the lowest performance among speciated $PM_{2.5,}$ except for nitrate.

4.3.2 Comparison of observed and simulated molecular markers at JST

Averages of measured and simulated organic molecular markers and performance of CMAQ simulation for each species (normalized mean fractional biases [MFB] and errors [MFE]) are given in Table 4.4. Normal alkanes (n-alkanes) concentrations are, in general, under-predicted (Figure 4.1), likely due to the missing vegetative detritus category in CMAQ. Vegetative detritus has the highest ratio of n-alkanes to organic carbon (OC) among all source categories. Nonacosane, octacosane, hexacosane, and pentacosane concentrations are better simulated in January, largely due to increased $PM_{2.5}$ emissions in natural gas combustion in the model simulation. Most n-alkanes have higher ambient levels in winter than in summer (Figure 4.2).

Measured levoglucosan and resin acids, except abietic acid (ABIET), are below their quantification limit in summer and increase in winter, similar to the seasonal variation found by Sheesley and Schauer (2004). Simulated levoglucosan matches well with measurements at JST in winter, while CMAQ still captures a significant amount of $PM_{2.5}$ coming from biomass burning even in summer with decreased biomass burning emissions owing to the open burning ban (Figure 4.3). Levoglucosan may decay in summer considering that potassium measurements, which is a tracer for biomass burning, do not go under the detection limits and there is no clear reason why biomass burning is completely diminished in summer.

Table 4.4 Performance of simulated organic molecular markers (July 2001)

Organic molecular markers	CMAQ mean (μgm^{-3})	Observation Mean (μgm^{-3})	MFB (%)	MFE (%)	Organic molecular markers	CMAQ Mean (μgm^{-3})	Observation Mean (μgm^{-3})	MFB (%)	MFE (%)
17a(H)-21b(H)-29-norhopane	0.31	0.18	49	57	tetracosanoic acid	0.03	1.19	-187	187
7h-benz(de)anthracen-7-one	0.21	0.26	13	65	pentacosanoic acid	0.00	0.16	-180	197
20R, aaa-cholestane	0.10	0.10	15	54	hexacosanoic acid	0.02	0.63	-185	185
20R+S, abb-cholestane	0.14	0.06	91	95	heptacosanoic acid	0.00	0.17	-147	197
abietic acid	0.71	0.38	122	139	octacosanoic acid	0.01	0.90	-193	193
anteiso-triacontane	0.11	0.35	-62	102	triacontanoic acid	0.01	0.98	-181	198
aluminum	40.80	50.57	14	53	17a(H)-21b(H)-hopane	0.34	0.24	36	49
benzo(b)fluoranthene	0.66	0.47	41	57	iso nonacosane	0.03	0.53	-157	174
benzo(e)pyrene	0.29	0.46	-25	63	iso-hentriacontane	0.25	0.59	-55	74
benzo(ghi)perylene	0.01	0.62	-160	193	indeno(cd)fluoranthene	0.02	0.33	-103	169
benzo(k)fluroanthene	0.90	0.33	96	96	indeno(cd)pyrene	0.08	0.66	-104	138
nonanal	5.50	4.19	58	81	levoglucosan	15.05	6.70	153	180
cholesterol	0.79	0.21	148	158	pentacosane	0.81	1.44	-37	55
20R+S, abb-ergostane	0.13	0.07	74	87	hexacosane	1.00	0.89	35	69
hexadecenoic acid	2.32	1.37	85	104	heptacosane	0.77	0.39	78	85
octadecenoic acid	27.98	1.67	181	181	octacosane	0.40	0.74	-47	55
tetradecanoic acid	0.28	1.13	-92	96	nonacosane	0.51	1.90	-107	107
pentadecanoic acid	0.07	0.58	-140	140	triacontane	0.07	0.52	-140	140
hexadecanoic acid	2.07	0.37	138	138	hentriacontane	0.69	2.15	-88	88
heptadecanoic acid	0.31	0.56	-41	61	odotriacontane	0.08	0.58	-135	135
octadecanoic acid	3.96	5.01	3	62	tritriacontane	0.22	0.79	-97	97
nonadecanoic acid	0.07	0.18	-80	82	tetracontane	0.00	0.47	-183	199
eicosanoic acid	0.21	0.70	-97	99	propionylsyringol	0.21	0.06	151	160
heneicosanoic acid	0.02	0.16	-143	159	silicon	54.79	116.36	-31	70
docosanoic acid	0.06	0.71	-163	163	20R+S, abb-sitostane	0.13	0.14	2	54
tricosanoic acid	0.02	0.30	-157	157	17a(H)-21b(H)-hopane	0.12	0.06	70	79

Table 4.5 Performance of simulated organic molecular markers (January 2002)

marker_id	CMAQ Mean (μgm^{-3})	Observation Mean (μgm^{-3})	MFB (%)	MFE (%)	marker_id	CMAQ Mean (μgm^{-3})	Observation Mean (μgm^{-3})	MFB (%)	MFE (%)
17a(H)-21b(H)-29-norhopane	0.29	1.40	-118	119	hexacosanoic acid	0.13	4.02	-185	185
7h-benz(de)anthracen-7-one	2.13	11.36	-111	115	heptacosanoic acid	0.00	0.56	-199	199
20R, aaa-cholestane	0.09	0.40	-111	113	triacontanoic acid	0.09	1.67	-173	173
20R+S, abb-cholestane	0.13	0.32	-71	78	triacontanoic acid	0.03	1.82	-188	188
abietic acid	19.81	0.81	184	184	17a(H)-21b(H)-hopane	0.32	1.36	-113	114
anteiso-triacontane	0.14	0.99	-146	146	iso nonacosane	0.04	2.01	-189	189
aluminum	64.55	11.81	141	141	iso-hentriacontane	0.30	1.38	-115	118
benzo(b)fluoranthene	6.75	1.78	120	120	indeno(cd)fluoranthene	0.08	0.46	-117	120
benzo(e)pyrene	2.83	1.56	67	73	indeno(cd)pyrene	0.35	1.41	-100	105
benzo(ghi)perylene	0.19	2.35	-156	156	isopimaric acid	4.17	2.16	86	96
benzo(k)fluroanthene	9.23	1.11	158	158	levoglucosan	406.53	399.33	12	56
nonanal	7.21	1.52	133	133	pentacosane	2.04	4.15	-54	69
cholesterol	1.04	0.89	16	43	hexacosane	1.51	2.97	-52	64
20R+S, abb-ergostane	0.12	0.36	-83	88	heptacosane	1.44	0.33	124	126
hexadecenoic acid	3.51	0.55	149	149	octacosane	0.68	1.72	-72	76
octadecenoic acid	48.80	4.58	169	169	nonacosane	1.60	3.35	-57	62
tetradecanoic acid	1.54	2.36	-45	52	triacontane	0.22	1.35	-133	133
pentadecanoic acid	0.38	1.06	-96	97	hentriacontane	1.17	3.94	-86	90
hexadecanoic acid	8.75	0.54	172	172	odotriacontane	0.12	1.31	-157	157
heptadecanoic acid	0.57	1.05	-49	62	tritriacontane	0.28	1.50	-125	125
octadecanoic acid	5.84	12.64	-60	68	tetracontane	0.00	0.89	-200	200
nonadecanoic acid	0.07	0.41	-134	134	8,15-pimaradien-18-oic acid	5.75	0.45	171	171
eicosanoic acid	0.22	1.74	-147	147	pimaric acid	2.07	1.92	28	57
heneicosanoic acid	0.02	0.88	-189	189	propionylsyringol	5.85	9.25	10	84
docosanoic acid	0.07	3.73	-191	191	sandaracopimaric acid	0.97	0.52	73	82
tricosanoic acid	0.03	1.65	-191	191	silicon	83.84	42.90	67	67
tetracosanoic acid	0.04	7.63	-198	198	20R+S, abb-sitostane	0.11	0.67	-130	130
pentacosanoic acid	0.00	0.84	-199	199	17a(H)-21b(H)-hopane	0.11	0.64	-136	136

Observed levels of levoglucosan increase from about 30ngm^{-3} (summer) to about 300ngm^{-3} (winter) (Figure 4.3). Increased wood combustion, especially for residential heating in winter, leads to increases in levoglucosan and resin acids in winter. Since relative increases of those species from summer to winter are greater in the urban and suburban areas than in the rural areas, urban areas experience a greater relative increase (Table 4.5).

Relative increases in the resin acids in winter were smaller than increases in levoglucosan levels, possibly due to differences in the types of wood that are burned in different seasons, and the types of combustion processes (i.e., fireplaces versus open burning). Resin acids mainly come from pine wood (Rogge et al., 1998), while levoglucosan is created by degradation of cellulose from both softwood and hardwood (Simoneit et al., 2000).

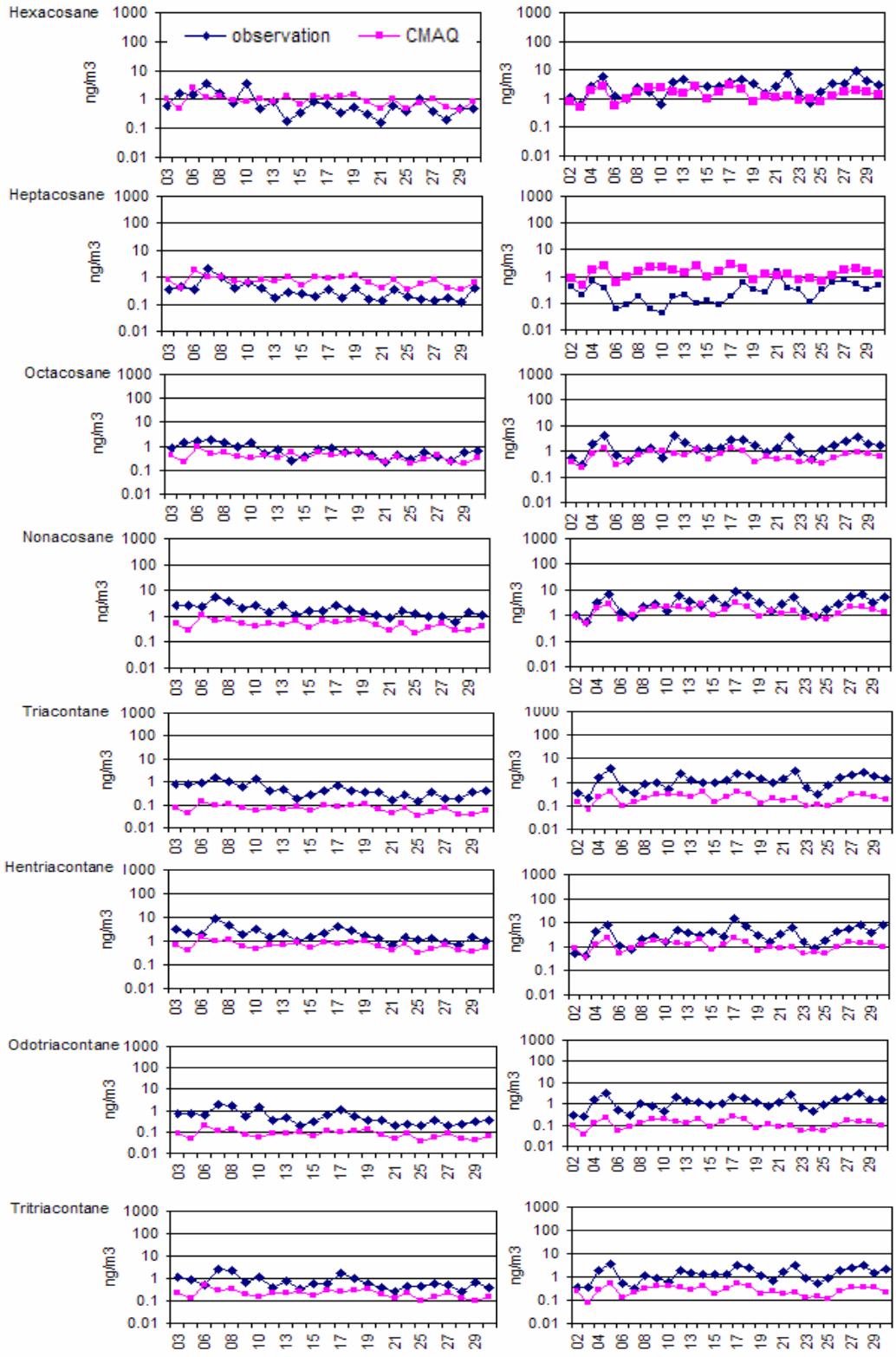


Fig. 4.1 Daily observations and CMAQ simulations of n-alkanes in July 2001 and January 2002, at Jefferson St., Atlanta.

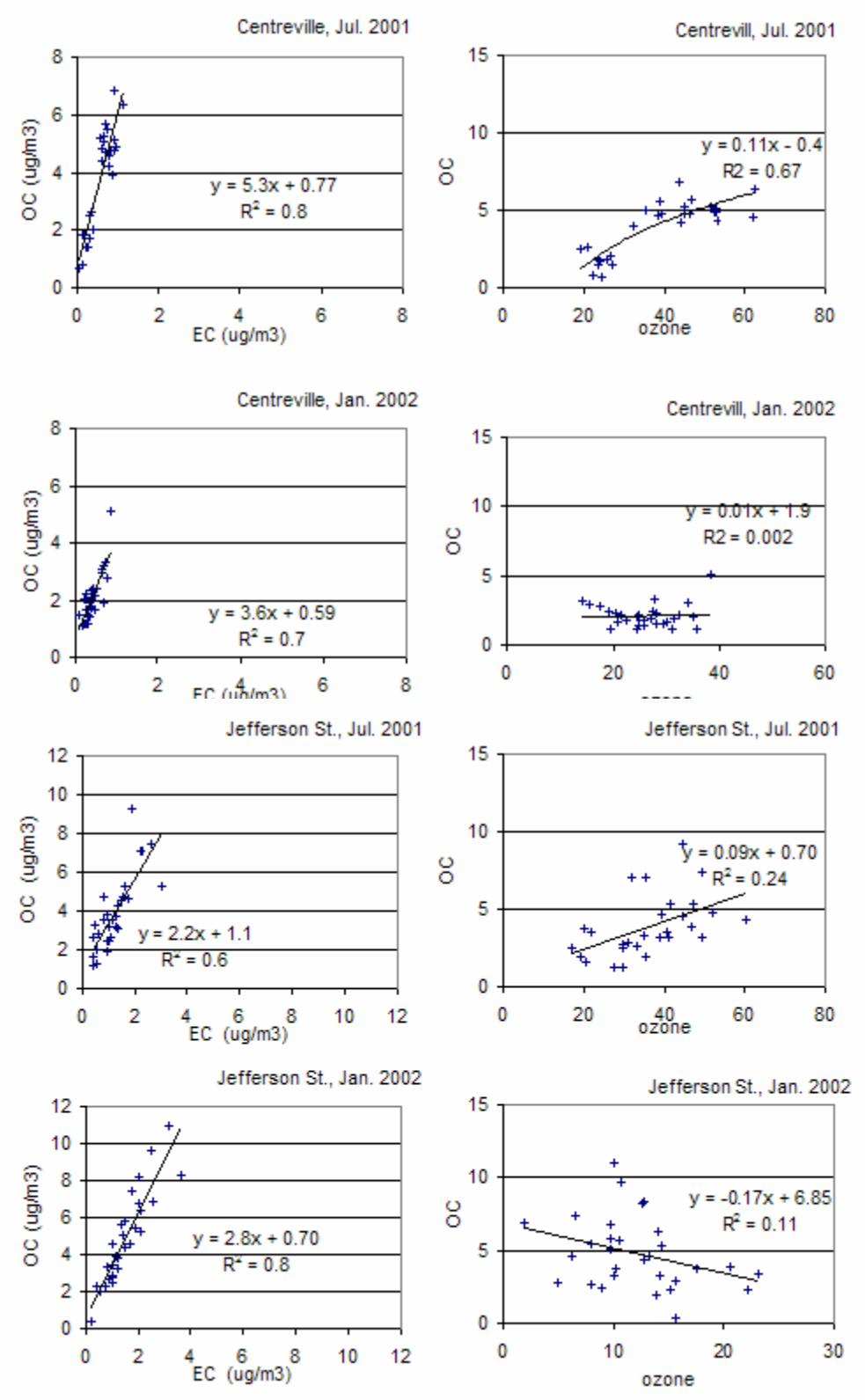


Figure 4.2 Scatter graph of ratios of measured OC vs. EC (left panel) an d OC with ozone (right panel)

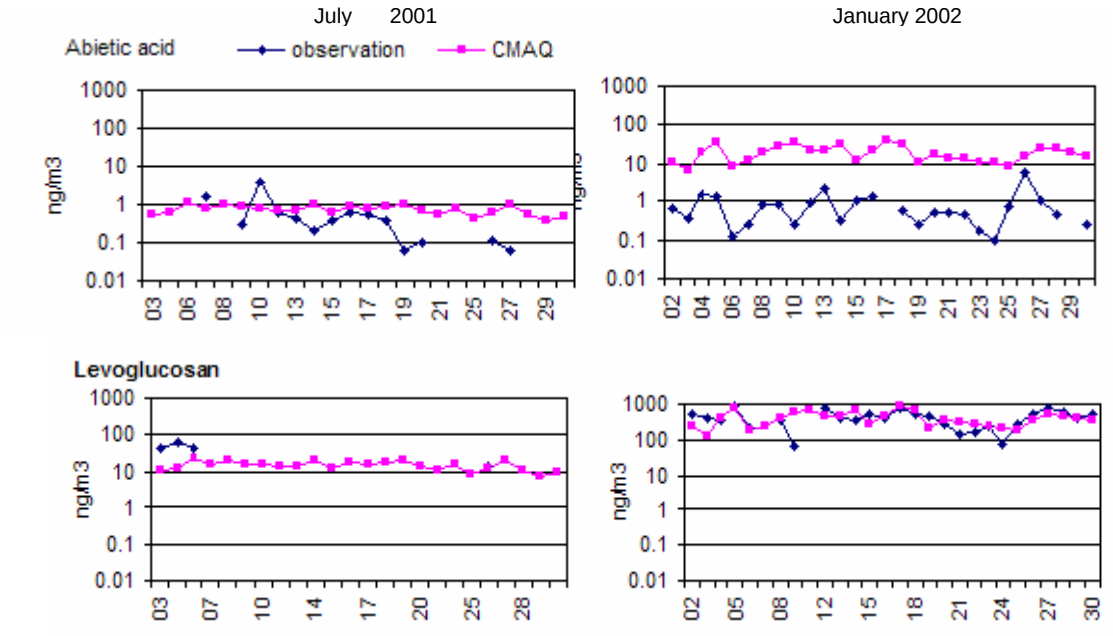


Figure 4.3 Observed and simulated of levoglucosan and resin acids in J uly 2001 and January 2002.

Table 4.6 Ratio of levoglucosan and resin acids concentrations in January to those in July.

	CTR	BHM	YRK	JST	OAK	GFP	OLF#8	PNS
CMAQ	9.8	10.8	22.6	36.0	2.9	3.1	8.4	9.7
Levoglucosan	3.7	6.3	6.3	13.0	2.6	11.3	8.2	16.6
Resin acids	2.0	3.0	8.2	4.0	2.5	3.9	2.0	8.7

CMAQ; Ratios calculated using simulated concentration of forest fires and fireplaces/woodstoves in CMAQ.

Levoglucosan, resin acids: Ratios calculated using measurements

Some hopanes and steranes, such as benzo(*de*)anthracen-7-one, 17α(*H*),21β(*H*)-29-norhopane and 17α(*H*)21β(*H*)-hopane, increase four to five-fold in winter, possibly due to increased emissions from cold starting engines (EPA, 2006). However, CMAQ exhibits little seasonal variation, possibly because it misses increased emissions in wintertime (Figure 4.4). Measurements at different ambient temperatures show that emissions of hydrocarbons during cold starts increased significantly at low temperatures. In the field, (Zielinska et al., 2004)) found increased emission rates of hopanes and steranes from gasoline vehicles as temperatures decreased. Increased condensationof SVOCs is another factor. Zhang et al. (Zhang and Wexler, 2004; Zhang et al., 2005) showed that decreased ambient temperatures lead to more nucleation and condensation of tail pipe emissions. However, increased condensation may not be significant in this study, since the high volume sampler used at the SEARCH stations has a positive artifact from organic gases (Turpin et al., 2000).

Model performance for PAHs varies markedly by species. Simulated concentrations of benzo(*b*)fluoranthene (BBF) and benzo(*e*)pyrene (BEP) are in good agreement with observations during summer although they are overpredicted in winter. Simulated benzo(*k*)fluoranthene (BKF) is higher than observations in both periods. Benzo(*ghi*)perylene (BGP), indeno(cd)pyrene (ICDP), and indeno(cd) fluoranthene (ICDF) levels are underpredicted (Figure 4.5). Based on source contributions estimated in CMAQ for each PAH, increases in BEP, BKF in winter are due to increased natural gas combustion. Most measured PAHs, except indeno(cd)fluoranthene (ICDF), increase in winter, but not as large as in simulations.

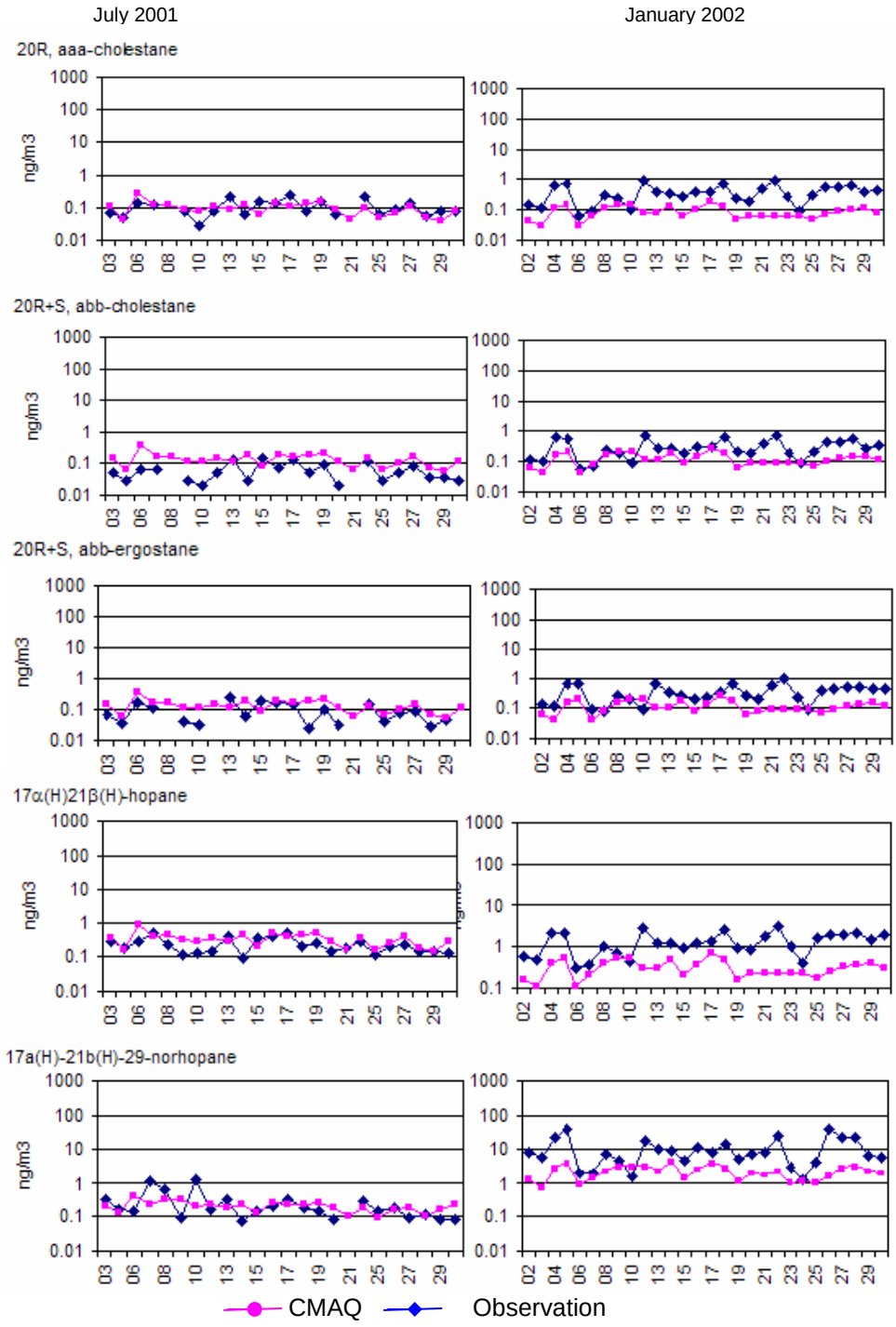


Figure 4.4 Daily observations and CMAQ simulations of hopanes and steranes in July 2001 and January 2002 at Jefferson St., Atlanta.

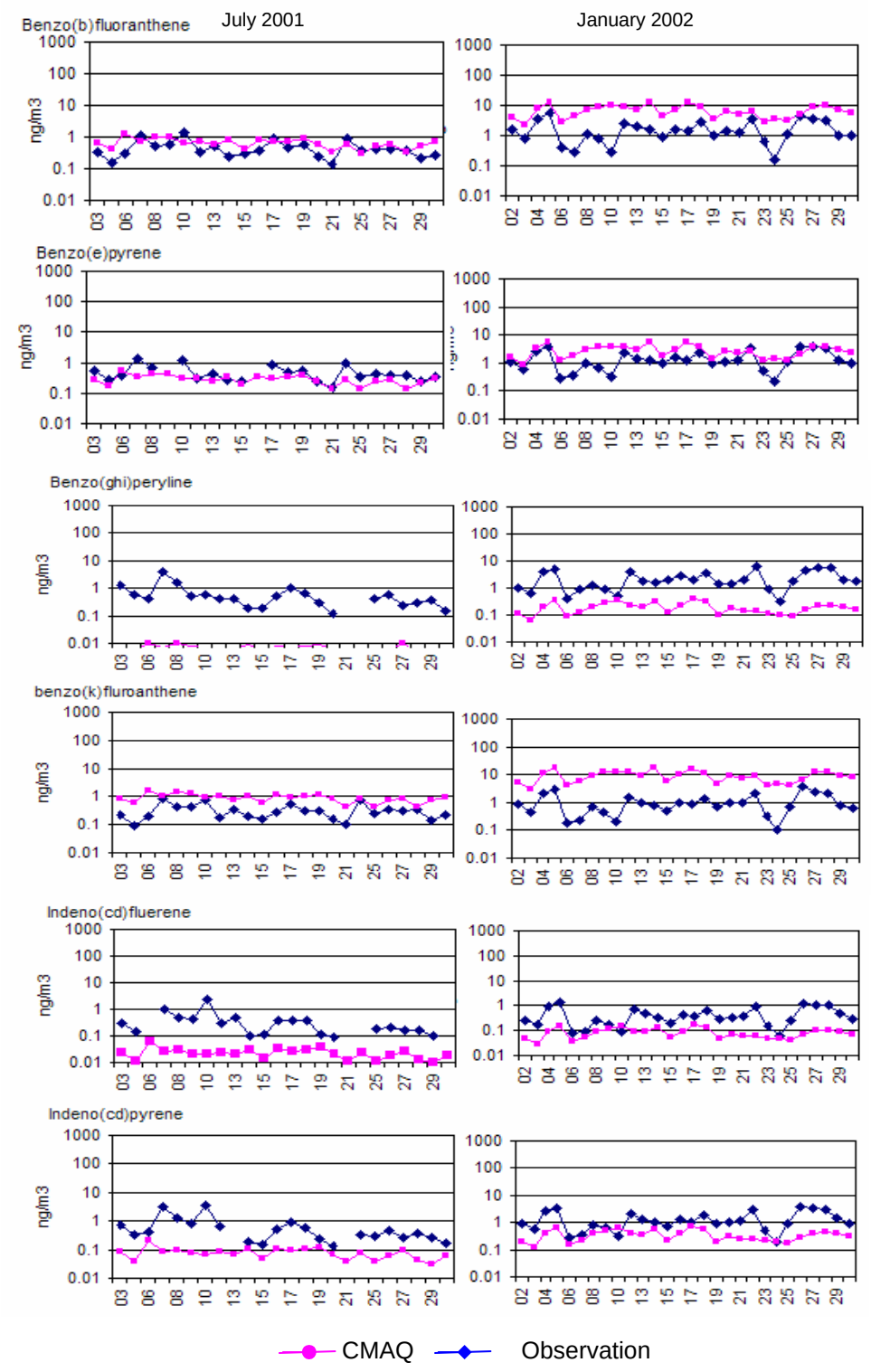


Figure 4.5 Observations and CMAQ simulations of PAHs in July 2001 and January at Jefferson St., Atlanta.

4.3.3 Scaling factors for emission estimates

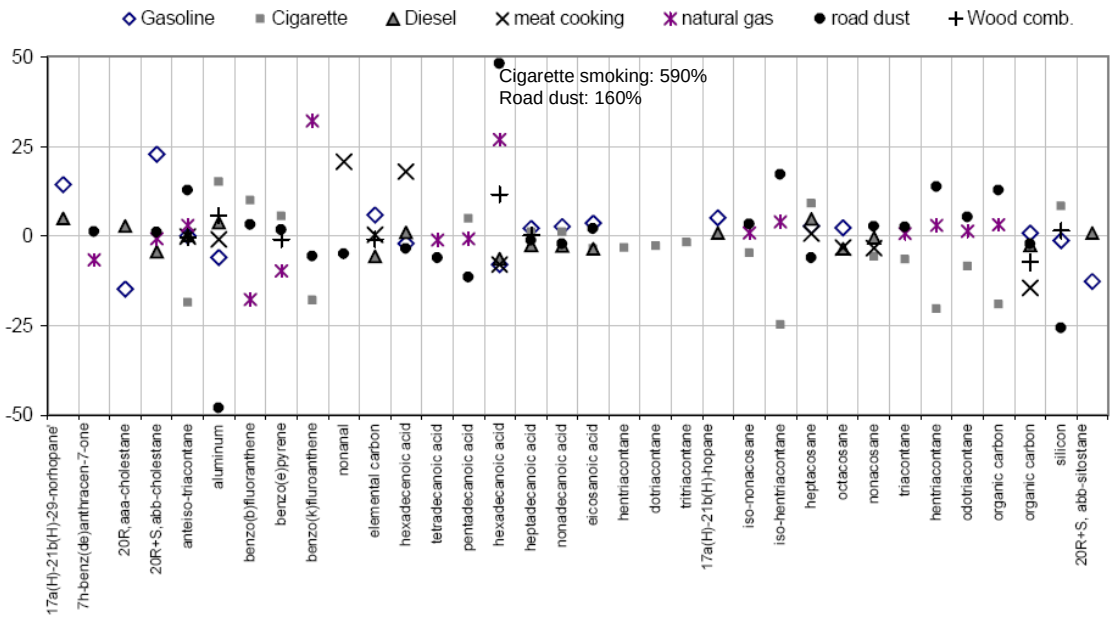
Scaling factors, whose values indicate likely levels of biases in the emission inventories, for each source category using all molecular markers ranged from 0.03 to 7.0, with some species having larger impacts on **F** than others. For example, EC is an important species for determining the scaling factor for gasoline ($f_{gasoline,JST}$) and diesel engine exhaust ($f_{diesel,JST}$), and Al and Si are important for the scaling factor for road dust ($f_{r.dust,JST}$). Sensitivity analysis of scaling factors at JST suggests that some species, such as ABIET and hexadecanoic acid (HDA), have unexpectedly large impacts on calculated scaling factors (Figure 4.6, a and b using SP_{Rogge}; c and d using SP_{Lough}). ABIET significantly lowers the scaling factor for wood combustion ($f_{wood,JST}$) by 150% in January (i.e., $f_{wood,JST}$ = 0.3 with ABEIT considered and 0.7 without). HDA decreases the scaling factor for natural gas combustion ($f_{n.gas,JST}$) by 80% (January), cigarette smoking by 500% (July), and road dust by 150%. This amount of change is unexpected, since more than 40 species are used in the regression analysis. The unanticipated large impact of one species may come from errors in measurements, uncertainties in source profiles, or other unknown factors. For example, large impacts of heptacosane (normal alkane, NA27) on most sources are observed only in winter, and it may come from measurement error. NA27 is less correlated with other normal alkanes in summer than in winter, while all the other normal alkanes are highly correlated with each other in both seasons (Table 4.7). On the other hand, species that have impacts larger than 100% in both seasons, such as HDA, ABIET, and octadecanoic acid (ODA) likely have biases in source profiles. Therefore, scaling factors were recalculated excluding ABIET, ODA, and HDA in summer, and NA27 additionally in winter.

Table 4.7 Correlation coefficients between measured normal alkanes in July 2001 and January 2002 at Jefferson St., Atlanta.

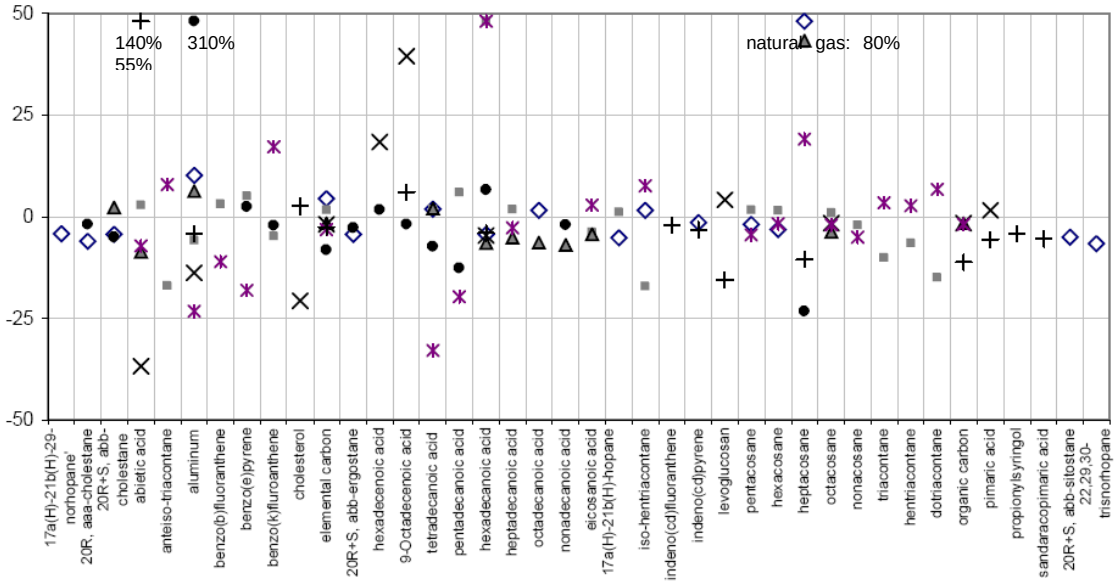
July	pentacosane	hexacosane	heptacosane	octacosane	nonacosane	triacontane	hentriacontane	dotriacontane	tritriacontane
hexacosane	0.89	-	-	-	-	-	-	-	-
heptacosane	0.90	0.80	-	-	-	-	-	-	-
octacosane	0.77	0.87	0.76	-	-	-	-	-	-
nonacosane	0.83	0.79	0.87	0.86	-	-	-	-	-
triacontane	0.79	0.90	0.79	0.95	0.88	-	-	-	-
hentriacontane	0.85	0.74	0.89	0.76	0.95	0.82	-	-	-
dotriacontane	0.82	0.83	0.85	0.84	0.90	0.93	0.92	-	-
tritriacontane	0.74	0.66	0.79	0.72	0.88	0.80	0.94	0.94	-
tetracontaine	0.89	0.75	0.93	0.70	0.87	0.74	0.95	0.85	0.84

January	pentacosane	hexacosane	heptacosane	octacosane	nonacosane	triacontane	hentriacontane	dotriacontane	tritriacontane
hexacosane	0.97	-	-	-	-	-	-	-	-
heptacosane	0.29	0.27	-	-	-	-	-	-	-
octacosane	0.85	0.87	0.23	-	-	-	-	-	-
nonacosane	0.72	0.76	0.20	0.85	-	-	-	-	-
triacontane	0.83	0.87	0.29	0.97	0.86	-	-	-	-
hentriacontane	0.62	0.66	0.19	0.72	0.94	0.77	-	-	-
dotriacontane	0.87	0.91	0.33	0.96	0.85	0.97	0.76	-	-
tritriacontane	0.84	0.87	0.36	0.90	0.90	0.94	0.86	0.97	-
tetracontaine	0.80	0.80	0.47	0.80	0.62	0.83	0.60	0.86	0.84

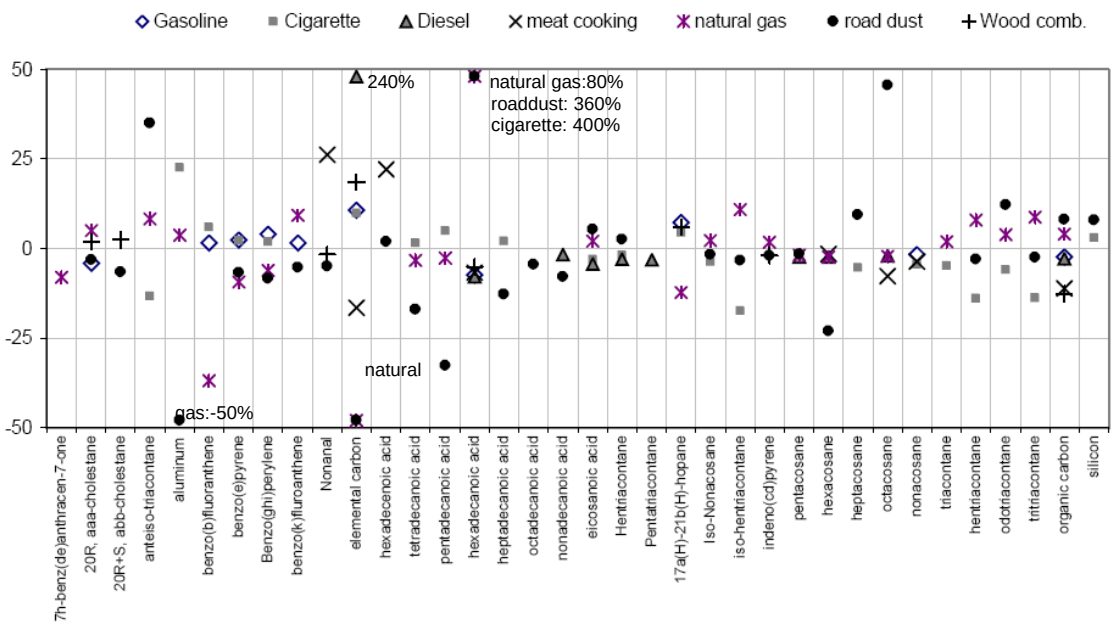
a) July 2001. SP_{Rogge}



b) January 2002. SP_{Rogge}



c) July 2001. SP_{Lough}



d) January 2002. SP_{Lough}

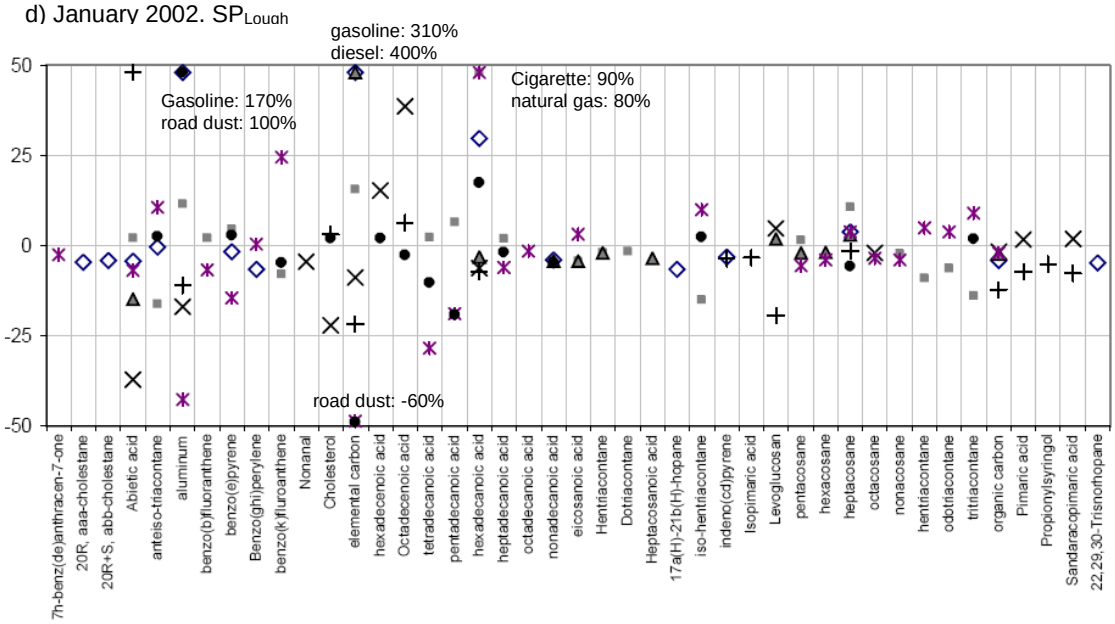


Figure 4.6 Normalized changes in scaling factors at JST a) in July 2001 using SPRogge, b) in January 2002 using SP_{Rogge}, c) July 2001 using SP_{Lough} and d) January 2002 using SP_{Lough}. Values in Y axis for "aluminum" are normalized changes in scaling factors when aluminum is excluded from a ridge regression analysis. Normalized changes less than ±1% are not shown in plots.

Scaling factors for JST and other SEARCH monitoring sites (Figure 4.7) suggest that missions from gasoline vehicles are estimated to be low by a factor of three in winter given the increased ambient levels of hopanes and steranes. Increases in emissions from diesel engines by 50% are suggested in both seasons, driven largely by EC since it is simulated low in CMAQ. Scaling factors for mobile sources estimated using trace metals in Chapter 4 range from 1.0 to 1.5 at JST, less than $F_{Lough,JST}$ or $F_{Rogge,JST}$. This is mainly because of larger EC to OC ratios in the vehicle source profiles in CMB studies using trace metals. The EC to OC ratios in LGO study are 0.4 (gasoline vehicles) and 3.7 (diesel vehicles), while they are 0.3 (gasoline vehicles) and 2.6 (diesel vehicles) in MM.

Using SP_{Lough}, as opposed to SP_{Rogge}, leads to larger scaling factors for gasoline vehicles in summer and the similar number in winter, because SP_{Lough} has lower ratios of hopanes and steranes to OC (summer) and a higher ratio of EC to OC (winter). Levels of hopanes and steranes are well captured using SP_{Rogge} in July (Figure. 4.4), and those species are the major contributors to $f_{Rogge,gasoline,JST}$ (Figure 4.6). The influence of hopanes and steranes on scaling factors decreases by switching SP_{Rogge} with SP_{Lough}. SP_{Lough} enlarges the gap between simulated and observed hopanes and steranes, which in turn increases the scaling factors. On the other hand, $f_{Rogge,gasoline,JST}$ is less dependent on hopanes and steranes in winter due to doubled concentrations of hopanes and steranes than in summer, leading to larger estimates in summer than in winter. This result suggests that the influence of species on scaling factors varies as seasons change altering the results of regression analysis. Thus, source profiles for mobile sources may differ by season (i.e., smaller ratios of hopanes and steranes in summer than in winter).

Sources other than vehicle exhaust, such as coal combustion, might have increased

ambient hopanes and steranes in winter, though a specific reason for this is not obvious. Increased contribution of coal burning is suspected based on an observed 22,29,30-trisnoneohopane to 22,29,30-trisnohopane ratio. The ratio at JST changed from 1.0 (summer) to 0.68 (winter), which indicates that the contribution of coal burning to hopanes increases in winter (Schnelle-Kreis et al., 2005).

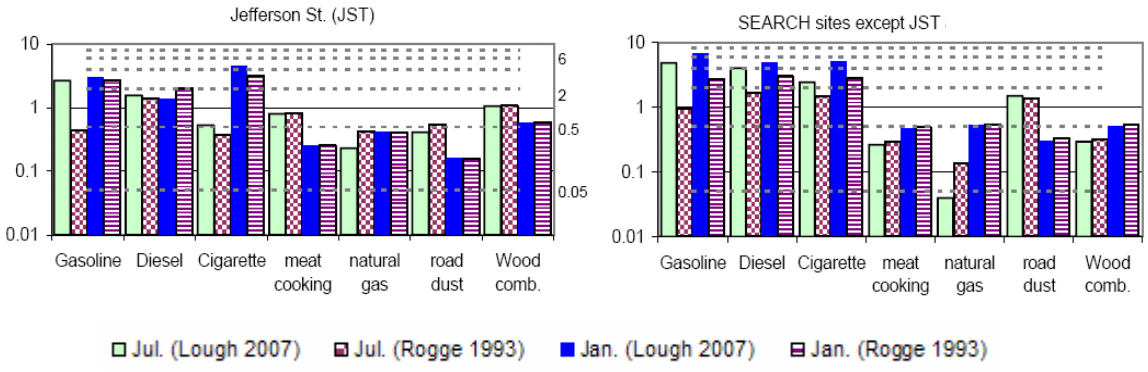


Figure 4.7 a) Scaling factors for emission impacts at Jefferson St. (JST) a nd the other SEARCH monitoring sites. “Rogge 1993”; using SP_{Rogge}. “L ough 2007”, SP_{Lough} b) Scaling factors are calculated using all the other S EARCH monitoring sites except JST.

Table 4.8 Scaling factors (± 95% confidence interval) for emission estimates at Jefferson St. (JST). “Rogge”; using SP_{Rogge}. “Lough”; using SP_{Lough}

	Jul. (Lough)	Jul. (Rogge)	Jan. (Lough)	Jan. (Rogge)
Gasoline exhaust	2.6 ± 0.22	0.4 ± 1.15	3.2 ± 0.03	2.7 ± 0.16
Diesel exhaust	1.6 ± 0.83	1.4 ± 0.82	1.5 ± 0.11	2.0 ± 0.11
Cigarette smoking	0.5 ± 0.34	0.4 ± 0.56	4.5 ± 0.17	3.1 ± 0.18
meat cooking	0.8 ± 0.47	0.8 ± 0.46	0.2 ± 1.22	0.3 ± 1.21
natural gas combustion	0.2 ± 1.15	0.4 ± 1.14	0.3 ± 1.34	0.4 ± 1.33
Road dust	0.4 ± 0.37	0.5 ± 0.36	0.1 ± 0.90	0.2 ± 0.89
Wood combustion	1.0 ± 0.02	1.1 ± 0.02	0.6 ± 0.67	0.6 ± 0.68
Biogenic SOA	1.0 ± 0.2	1.1 ± 0.2	0.9 ± 0.05	0.9 ± 0.05
Anthropogenic SOA	1.0 ± 0.01	1.0 ± 0.01	1.0 ± 0.01	1.0 ± 0.01

Overestimates in source impacts of wood combustion in winter appear larger than in summer, with $f_{wood,JST}$ around 0.5 in winter and 0.* in summer. This result is consistent with what has been found in the previous study using trace metals. Considering the overestimated OC in winter at JST (NMFB 47%), a decrease in emissions from wood combustion seems reasonable, which further suggests possible biases in levoglucosan speciation in the wood combustion source profile. Simulated levoglucosan matches very well with observed values with a NMFB of 12%, while ridge regression analysis supports that wood combustion emissions should be decreased by 50%. This implies that the ratio of levoglucosan to OC defined in the source profile (12%) is low and it may be increased to about 20%. This number falls into the range of levoglucosan ratios found in previous studies in which it was found that levoglucosan emission rates differed by wood types, regions and boiler types, with fractions of levoglucosan ranging from 0.3% (Hedberg et al., 2006) to 30% (Schauer et al., 2001). The wood combustion source profile used here was derived from measurements in Southern California using three wood types found in the southeast (Rogge et al., 1998), but differ from the actual emission characteristics of fireplaces and woodstoves. Given the wide range of emission rates of levoglucosan, it is questionable whether or not levoglucosan is suitable as a tracer of wood combustion (Hedberg et al., 2006). A large uncertainty in the levoglucosan fraction in receptor modeling should be considered.

Results suggest emissions from natural gas combustion should be decreased by 60~80% in both seasons, driven by overestimated concentrations of BKF, BBF and BEP. Emissions from meat cooking and road dust appear to be biased high but differ markedly by season. $f_{meat,JST}$ is 0.8 in summer, 0.3 in winter, and $f_{r.dust,JST}$ is 0.5 in summer and 0.2 in

winter.

Overall, F_{others} have a similar trend with F_{JST}, which suggests increases in emissions from mobile sources and decreases in wood combustion, meat cooking, and road dust. $f_{gasoline,others}$ has a similar seasonal change to $f_{gasoline,JST}$, but is larger, ranging from 1.0 to 6.0. This is because EC is simulated significantly low in $Other_{SEARCH}$. (NMFB -70% in July and -35% in January). $f_{gasoline,others}$ using SP_{Rogge} is less than 1.0 in summer, as is $f_{gasoline,JST}$, due to the same reason: hopanes and steranes are simulated high. $f_{cigarette,others}$ is larger than $f_{cigarette,JST}$ mainly because of under-predicted levels of normal alkanes in $Other_{SEARCH}$. Cigarette smoking has the second largest ratio of normal alkanes to organic carbon, following vegetative detritus. Vegetative detritus is not included in the CMAQ simulation; thus, an increase in cigarette smoking emissions is suggested to compensate for the lack of vegetative detritus. Cigarette smoking emits other organic molecular markers as well, such as EC, fatty acids, levoglucosan, and anteiso-triacontane, but the amounts of those species are so low that normal alkanes are the major factors determining the value of $f_{cigarette,others}$. This may suggest that vegetative detritus needs to be considered in CMAQ simulations to adequately simulate normal alkanes. $f_{wood,others}$ is about 0.5 due to over-predicted PAHs and resin acids, such as ICDP, isopimaric acid (IPIMA), pimaric acid (PIMA), propionylsyringol (PSYR), and sandaracopimaric acid (SAND). Road dust increased due to under-predicted Al and Si in July.

4.3.4 Performance of adjusted $PM_{2.5}$ levels

Calculated $PM_{2.5}$ source contributions at JST and $Other_{SEARCH}$ were updated with the estimated scaling factors, to assess improvement in $PM_{2.5}$ simulations. For example, $PM_{2.5}$ impacts from the seven source categories at JST are multiplied with corresponding scaling factors, and the sums of these updated values are compared with observed PM2.5. Changes in MFB and MFE of $PM_{2.5,}$ OC and EC using a different set of F ($F_{Lough,JST}$, $F_{Rogge,JST}$, $F_{Lough,others}$ and $F_{Rogge,others}$) largely decreases NMFB and NMFE (Figure 4.8). NMFB and NMFE of total $PM_{2.5}$ increased with $F_{Lough,JST}$ in July, mainly because of the large scaling factors for mobile sources. Overall, improvement at JST was better than at other sites. In Chapter 4, the scaling factors vary significantly between different areas (more than 50% difference between urban, rural, and suburban areas); thus, using one scaling factor for different regions limits improvement in $PM_{2.5}$ simulations. Different measurement intervals may be another reason, since there are six sets of monthly molecular marker observations for the $Other_{SEARCH}$ sites, while 25 sets of daily values per month are available at JST.

NMFE and NMFB of OC at other sites improved less in summer than in winter primarily because more secondary organic carbon (SOC) is formed in summer. If there is more SOC, it will lower the degree of improvement in primary OC simulations. However, improvements in NMFB and NMFE for OC at JST in both seasons are similar, suggesting that the contribution of SOC at JST may vary less with the season than other SEARCH sites, partly due to the increased primary emissions in areas around JST. Measured OC, EC, and ozone concentrations are compared to each other in order to evaluate seasonal changes in the contribution of SOC. Larger OC/EC is usually related to

increased SOC while smaller values are more influenced by $PM_{2.5}$ primary sources (Chu, 2005). At CTR (located in a rural area), the OC/EC ratio in summer is larger (5.3) than in winter (3.6) (Figure 4.10), and ozone and OC concentrations are strongly correlated in summer (R^2=0.67), but not in winter (R^2=0.0). On the contrary, OC/EC at JST does not vary significantly with season (2.2 in summer and 2.8 in winter) and there is weak correlation between OC and ozone in both seasons (R^2=0.2 in summer and R^2=0.1 in winter).

Changes in performance after applying scaling factors that are calculated with trace metals and with organic molecular markers are comparable, but with important differences (Figure 4.9). NMFEs of updated $PM_{2.5}$ using different species are quite similar to each other, but NMFEs and NMFBs of OC, and EC are different. In July 2001, a set of source profiles used in the CMB regular model in Chapter 4 (Lee et al., 2008) achieved the greatest improvement in OC and EC (except OC in January 2002). Meat cooking and natural gas combustion categories in MM cover more primary OC sources than SP_{RG} (20% more based on CMAQ source apportionment results), which is expected to improve OC simulations. Much less improvement in OC simulation using SP_{Lough} in summer indicates that 1) the use of daily observations in RG gives more improvement than the use of monthly data in MM at $Other_{SEARCH,}$ in terms of assessing emission inventories using regression analysis and/or 2) meat cooking and natural gas combustion source profiles may have significant impact on OC emissions.

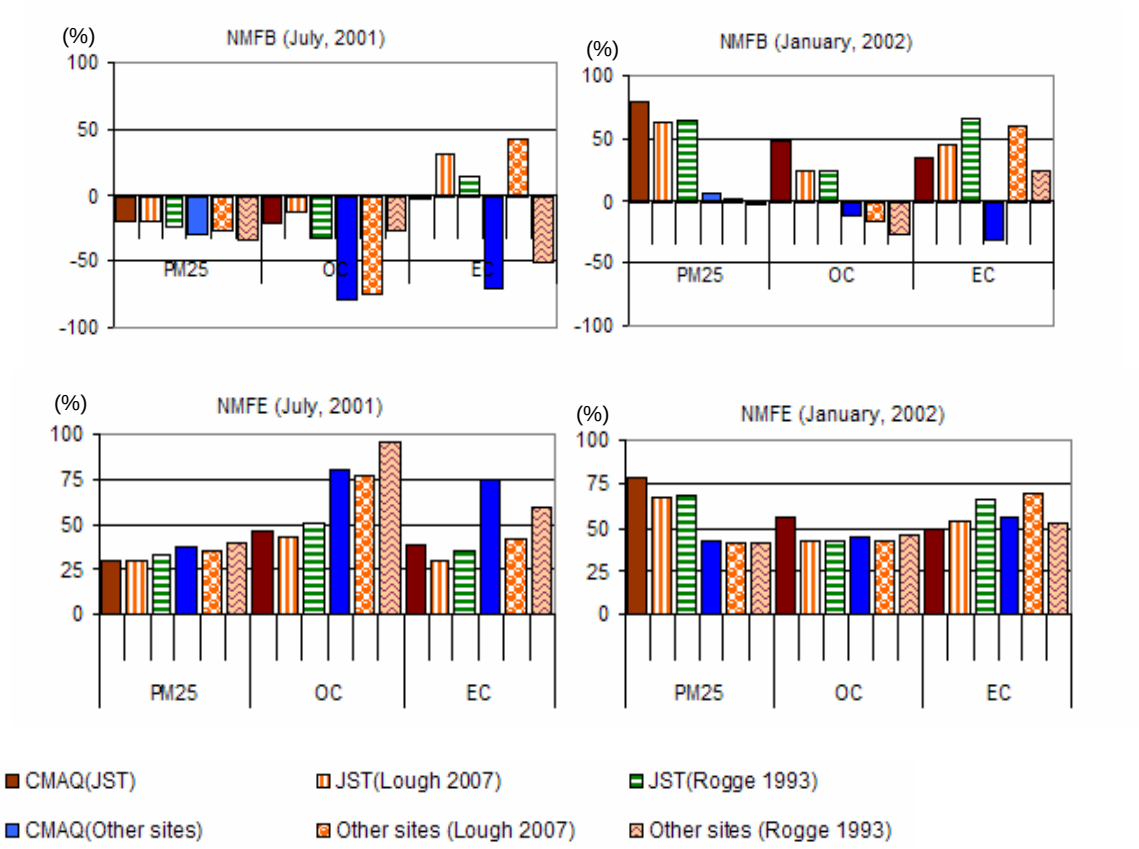


Figure 4.8 Changes at JST and in $Other_{SEARCH}$ of normalized mean fractiona l bias (NMFB) and error (NMFE). Each symbol represents 1) CMAQ (JST) : performance of original CMAQ simulation at Jefferson St., 2) JST (Lough 2007): performance of updated $PM_{2.5}$ simulations using scaling factors ($F_{Lough,JST}$) at JST, 3) JST (Rogge 1993): updated performance with a mobile source profile by Rogge et al. (1993) at JST ($F_{Rogge,\ JST}$), 4) CMAQ (Other sites): performance of original CMAQ simulation at other SEARCH monitoring sites (except JST) and 5) Other sit es (Lough 2007): performance of updated $PM_{2.5}$ simulations at other SEARCH monitoring sites (except JST) using F_{Lough}.

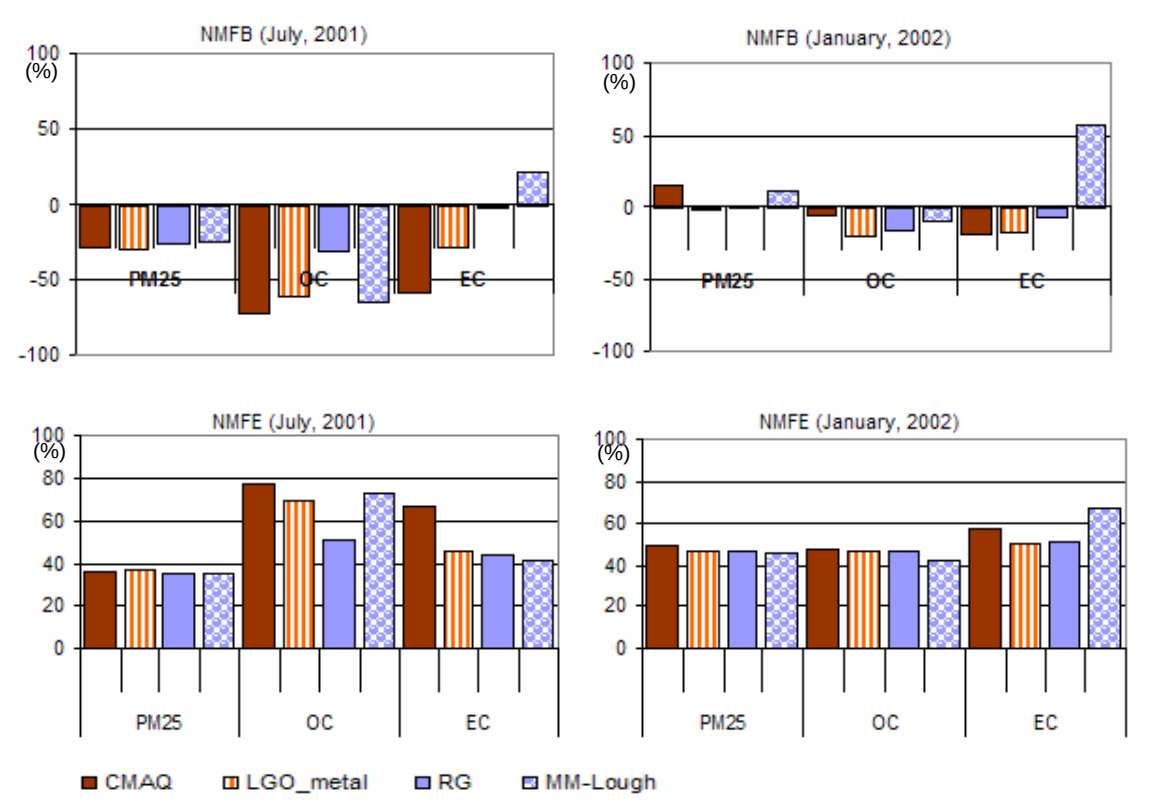


Figure 4.9 Normalized mean fractional bias (NMFB) and error (NMFB) at all SEARCH monitoring sites except PNS using different CMB source profiles. 1) CMAQ: performance of original CMAQ simulation, 2) LGO_metal: performance of updated $PM_{2.5}$ simulations using scaling factors calculated with source profiles from LGO study, 3) RG: updated performance with a mobile source profile in RG study, 4) MM-Lough: performance of updated $PM_{2.5}$ simulations using F_{Lough}.

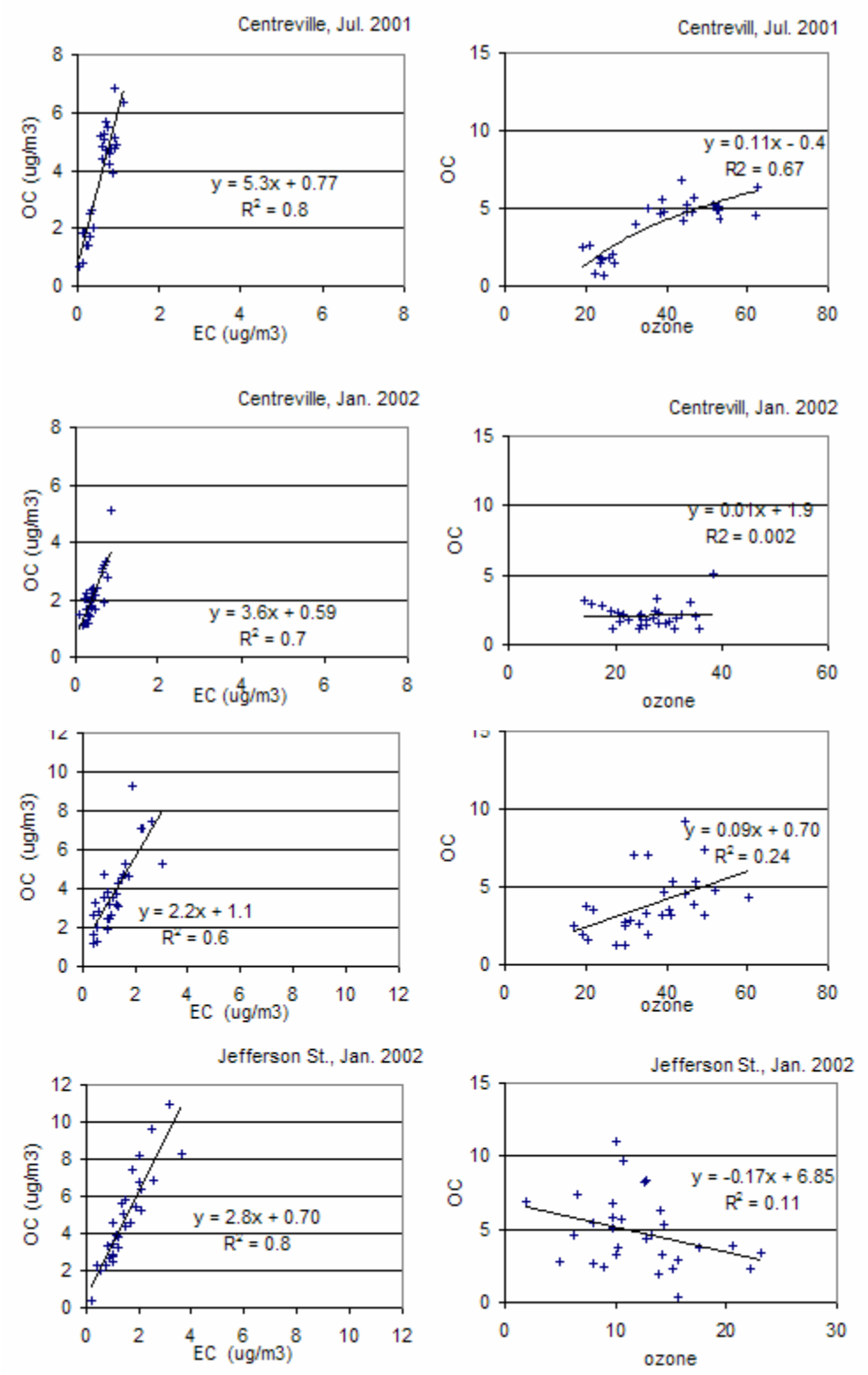


Figure 4.10 Scatter graph of ratios of measured OC vs. EC (left panel) and OC with ozone (right panel).

4.4. CONCLUSION

Emission estimates are a key component in air quality modeling, and a major source of uncertainty as well. Potential biases in the $PM_{2.5}$ emission inventories were quantified using regression analysis with organic molecular markers. Some molecular markers have noticeable seasonal variation that is not captured in CMAQ simulations, suggesting that the emission inventory does not capture temporal changes in source activities or in emission factors. The scaling factors for mobile sources using SP_{Lough} are most sensitive to EC while those using SP_{Rogge} respond to EC, hopanes and steranes. It may be necessary to use different mobile sources in different seasons due to strong seasonality in observed hopanes and steranes. Levoglucosan is an important tracer for wood combustion, and results suggest it is underestimated in the source profile. Overall, adjusting results using scaling factors improved model results in both seasons. However, the degree of improvement varied with the set of source profiles used and the number of measurements available, indicating that the selection of source profiles and the possession of substantial data are important in identifying emission inventory biases.

CHAPTER 5

MULTIGENERATIONAL SECONDARY ORGANIC AEROSOL

5.1 INTRODUCTION

Organic aerosol is a major component of fine particulate matter ($PM_{2.5,}$ PM with a diameter of less than 2.5μm) in most areas of the U.S. The importance of controlling organic aerosol in order to improve ambient air quality is growing as sulfate has been reduced significantly due to extensive controls (Evans et al., 2008). Organic aerosol is of either primary (POA) or secondary (SOA) origin. The major sources of POA are forest fires, biomass burning, meat cooking and mobile sources (Hopke et al., 2003; Kim et al., 2003; Schauer and Cass, 2000; Ying et al., 2004). Levels of POA tend to be higher in winter than in the summer due to increased wood burning and reduced dispersion. SOA is formed from anthropogenic and biogenic volatile organic carbons (VOCs), through photochemical reactions and other transformations. SOA tends to be elevated in summer due to increased emissions of biogenic precursors and photochemical reactions. Previous studies have shown that SOA may range from 20% to 70% of total organic aerosol depending on the season and location.

Chemical transport models (CTM) often simulate organic carbon (OC) lower than observations, possibly due to missing mechanisms and/or missing precursors of SOA. SOA formation from monoterpenes and aromatics are a major source of SOA in CTMs. New pathways or precursors related to SOA have been found in experimental studies in

recent years, including the increase of SOA yield and formation rate in acidic environments (Iinuma et al., 2004; Jang et al., 2002), oxidation of volatilized POA (Robinson et al., 2006; Robinson et al., 2007), polymerization and oligomerization (Jang et al., 2006; Jang et al., 2002; Kalberer, 2004), and SOA formation from isoprene (Claeys et al., 2004). Isoprene as a precursor of SOA is important as it may contribute significantly due its abundance (Griffin et al., 1999; Guenther et al., 1995; Guenther et al., 2006). Sesquiterpene emissions (e.g., Sakulyanontvittaya et al. (2008) may also be important progress as well, as the aerosol yield is much larger (between 20% and 70%) than those of monoterpenes and isoprene (7% and 2% respectively) (Griffin et al., 1999).

Recent studies suggest there is additional SOA formation from oxidation of semi volatile VOCs (SVOCs). First-generation SVOCs from reactions of VOC precursors may react further, producing secondary or tertiary SVOCs (multigenerational SVOCs, MSVOCs), and form additional aerosol (Ng et al., 2006). Aerosol yields used in CTM, such as the Community Multiscale Air Quality (CMAQ), count only SOA partitioned from first-generation SVOCs, produced from gas phase precursor species.

Studies have been done in order to improve SOA simulation in the air quality model by adding isoprene SOA (van Donkelaar et al., 2007; Zhang et al., 2007) and sesquiterpene SOA (Pun et al., 2006; Sakulyanontvittaya et al., 2008), and using the detailed reactions of monoterpene species in the CB4 mechanism (Pun et al., 2006). Here, we focus on updating biogenic SOA and multigenerational aerosol formation from both biogenic and anthropogenic SVOCs. The SOA module in CMAQ version 4.5 with SAPRC99 is updated by 1) modifying SOA formation from monoterpenes using species-specific parameters and emissions, 2) adding additional SOA from isoprene and

sesquiterpene and 3) simulating multi-generational aerosol by photochemical oxidation (i.e., aerosol formed from second or subsequently generated SVOC products). Parameters related to SOA in CMAQ, such as enthalpy of vaporization, saturation vapor pressure, the mass stoichiometric coefficient, and sensitivities of SOA to these parameters, are studied to understand the possible biases in CMAQ BSOA simulations.

5.2 METHODS

5.2.1 SOA simulation in CMAQ

In CMAQ, VOCs are simulated to react with hydroxyl radical (OH), nitrate radical (NO_3) and ozone (O_3), producing SVOCs:

$$VOC_J + O_3 \rightarrow \alpha_{J,O_3,1} P_{J,O3,1} + ... + \alpha_{J,O_3,m} P_{J,O3,m} + \alpha_{J,SVOC} SVOC_J \quad (5\text{-}1)$$

$$VOC_J + OH \rightarrow \alpha_{J,OH,1} p_{J,OH,1} + ... + \alpha_{J,OH,n} P_{J,OH,n} + \alpha_{J,SVOC} SVOC_J \quad (5\text{-}2)$$

$$VOC_J + NO_3 \rightarrow \alpha_{J,NO3,1} p_{J,NO3,1} + ... + \alpha_{J,NO3,n} P_{J,NO3,n} + \alpha_{J,SVOC} SVOC_J \quad (5\text{-}3)$$

where $\alpha_{J,i,n}$ is a mass stoichiometric coefficient for the nth product ($P_{J,i,n}$) that is produced from the reaction of VOC species J (VOC_J) with an oxidant i. It is assumed that a precursor VOC_J produces $SVOC_J$ with similar properties regardless of oxidants. Some VOCs produce one SVOC, and others produce two SVOCs (high volatility and low volatility OCs). The SVOCs produced have a low enough vapor pressure to partition into the aerosol phase, i.e., $SVOC_J$ exists in both gas and condensed phases in the atmosphere. The total concentration of $SVOC_J$ ($C_{tot,SVOCJ}$) is a summation of $SVOC_J$ in the gas ($C_{gas,SVOCJ}$) and aerosol phases ($C_{aer,SVOCJ}$). $C_{gas,J}$ is equal to or less than saturation vapor

pressure of $SVOC_J$, $C_{sat,J}$. $C_{sat,J}$ is calculated from the gas/particle partitioning coefficient of organic matter (K_{om}) that is measured in chamber studies:

$$C_{sat,J} = \frac{760 \times R \times T}{MW \times \xi_J \times K_{om,J}} \quad (5\text{-}4)$$

where T is in Kelvin, ξ_J is the activity coefficient of species J. $C_{sat,J}$ is temperature-dependent with the enthalpy of vaporization ΔH. The reference temperature of C_{sat} (T_{ref}) is 298K in CMAQ and adjusted to given temperature T:

$$C_{sat,T} = C_{sat,Tref} \times \frac{T_{ref}}{T} \times \exp\left[\frac{\Delta H}{R}\left(\frac{1}{T_{ref}} - \frac{1}{T}\right)\right] \quad (5\text{-}5)$$

where ΔH is enthalpy of formation and drives the relationship between saturation vapor pressure and temperature. ΔH of 156KJ/mol, as measured by Tao and McMurry (1984) using an individual aerosol species, was used in CMAQ until version 4.7 when It was changed to 40kJ/mol for biogenic precursors and 18 kJ/mol or 40kJ/mol for anthropogenic precursors. Offenberg et al. (2006) measured effective enthalpy of vaporization (ΔH^{eff}) using the photochemically produced SOA from α-pinene (APIN), isoprene and toluene, which they suggest the atmospheric composition of SOA. ΔH^{eff} is about 40KJ/mol for APIN and 18kJ/mol for isoprene, and ranges from 11 to 15 kJ/mol for toluene. ΔH is set as 40kJ/mol in the base simulation (named as **ΔH40** case) for all precursors in this study.

5.2.2 Modifications to the SOA module

(1) Detailed monoterpenes species

In CMAQ, five major monoterpene species are lumped into one species (TRP). α_{TRP} and $C_{sat,TRP}$ defined in CMAQ ($\alpha_{TRP}^{CMAQ}, C_{sat,TRP}^{CMAQ}$) are a weighted average of $\alpha_{J,VOC}$ and $C_{sat,J}$ of five species using mass contributions from Griffin et al. (1999) :

$$C_{sat,TRP} = \frac{\sum_{i=1}^{N} \alpha_i /(1 + K_{om,i} \times M_0)}{\sum_{i=1}^{N} \alpha_i K_{om,i} /(1 + K_{om,i} \times M_0)} \quad \text{N = 1, 5} \quad \text{(5-6)}$$

$$\alpha_{TRP} = \sum_{i=1}^{N} \alpha_i \times Ratio_{i,weight} \quad \text{(5-7)}$$

Aerosol yield, a ratio of aerosol produced to ROG reacted, is calculated using $a_{J,VOC}$, K_{om} and pre-existing aerosol mass (M_0, $\mu g\ m^{-3}$) in the air:

$$Y_i = M_0 \sum \frac{\alpha_{J,SVOC} K_{om,J}}{1 + K_{om,J} M_0} = M_0 \sum \frac{\alpha_{J,SVOC}}{1/C_{sat,J} + M_0} \quad \text{(5-8)}$$

where R is the ideal gas constant.

Major monoterpene species include α-pinene (APIN), β-pinene (BPIN), Δ^3-carene (DCAR), Δ-Limonene (DLIM) and sabinene (SABN). In CMAQ, mass contributions of these species over the U.S. continent are set as 40%, 25%, 15%, 10% and 10%, respectively, based on Griffin et al. (1999). In The Biogenic Emissions Inventory System (BEIS) version 3.12, emissions from 14 monoterpene species are separately calculated, and their mass contributions are different from the previously defined composition. For example, APIN contributes around 50% of total monoterpene in the Southeast and from 25% to 30% in the Northeast. BPIN contributes 30% in parts of Georgia and Florida and

less than 20% in most of the U.S., while DCAR contributes less than 5% in the Southeast. Changes in composition alter α_{TRP} and $C_{sat,TRP}$. For example, α_{TRP} decreases as the ratio of APIN rises, while becoming higher as the ratio of DLIM increases. Eight of 14 monoterpene species have individually measured α and C_{sat} (Table 5.1)(Griffin et al., 1999). α_{TRP} and $C_{sat,TRP}$ are calculated based on eight species' mass contributions to total monoterpenes using the equation (7) and (8) with N = 8.

Table 5.1 Reaction rates, mass stochiometric coefficients (α), and Saturation vapor pressure (Csat) of precursors of secondary organic aerosol in the updated CMAQ.

Precursors (abbr.)		Oxida-nts	Reaction rate (moles s^{-1} cm^{-3})	no. SVOCs	SVOCs Properties (T_0=298K)	
					α	C_{sat} μgm^{-3}
ALK	Alkanes	OH	$1.11\times10^{-11}exp(-52/T)$	1	0.07	0.02
OLE	Alkenes	OH	$1.74\times10^{-11}exp(384/T)$	2	0.36	111.11
		O_3	$5.02\times10^{-16}exp(-461/T)$		0.32	1000.00
		NO_3	7.26×10^{-13}			
ARO1	Toluene	OH	$1.81\times10^{-12}exp(355/T)$	2	0.07	10.51
					0.14	292.88
ARO2	Xylene	OH	2.64×10^{-11}	2	0.04	13.26
					0.17	397.75
TRP1	Monoterpenes	OH	$1.83\times10^{-11}exp(449/T)$	2	0.09	6.00
		O_3	$1.08\times10^{-15}exp(-821/T)$		0.39	109.00
		NO_3	$3.66\times10^{-12}exp(-175/T)$			
ISOPRENE	Isoprene	OH	$2.50\times10^{-11}exp(408/T)$	2	0.03	0.62
		O_3	$7.86\times10^{-15}exp(-1912/T)$		0.23	116.01
		NO_3	$3.03\times10^{-12}exp(448/T)$			
New reactions						
CARYL	b-caryophillene	OH	2.00×10^{-10}	1	1	2.19
		O_3	1.16×10^{-14}			
		NO_3	1.90×10^{-11}			
HUMUL	a-humulene	OH	2.90×10^{-10}			
		O_3	1.17×10^{-14}			
		NO_3	3.50×10^{-11}			
OSSQT	Other	OH	6.80×10^{-11}			
	sesquiterpenes	O_3	6.21×10^{-17}			
		NO_3	8.29×10^{-12}			

(2) Additional SOA formation from isoprene and sesquiterpene

The SAPRC99 mechanism includes isoprene reactions with OH, O_3 and NO_3, thus no changes in gas-phase reactions are made. However, two SVOC products (one with low volatility [LSVOC] and the other with high volatility [HSVOC]) are added as products of isoprene reactions with ozone, hydroxide, and nitrate radicals using parameters from Kalberer et al.(2006) (Table 5.1). Emissions from three sesquiterpenes are estimated in MEGAN (β-caryphyllene [CARYL], α-humulene [HUMUL] and other more slowly reacting sesquiterpenes [OSQT]). Shue and Atkinson (1995) measured reaction rates of five sesquiterpenes with OH, O_3 and NO_3. CARYL and HUMUL reaction rates are taken directly from their study and OSQT reaction rates are an average of a-cedrene, a-coparene and longifolene (Shu and Atkinson, 1995) (Table 5.1). Alkanes and sesquiterpenes have one SVOC product and the others have two SVOC products with high volatility (HSVOC) and low volatility (LSVOC). Corresponding α and Csat with HSVOC and LSVOC are named as named as α_H and $C_{sat,H}$, and α_L and $C_{sat,L}$ for LSVOC in sequence.

5.2.3 SOA formation from multigenerational SVOCs

The exact chemical products of the n-generated process are not known and are hard to parameterize (Altieri et al., 2006; Jang et al., 2006; Kalberer, 2004). Here, aged organic aerosol (AOA) is assumed to be produced from reactions of SVOCs with oxidants (OH, NO_3 radical and ozone) with reaction coefficients *ks* taken from the precursors' reaction rates in the SAPRC99 chemical mechanism. It is assumed that HSVOCs react with OH , NO_3 or ozone, evolve into LSVOCs, and then LSVOCs into multi-generational SVOCs

(MSVOCs). If the precursor A produces only one SVOC, then products from the following SVOC reactions are assumed to be MSVOCs that condense directly (i.e., aged organic aerosol. AOA), and unlike the other SVOCs (first-generation precursor reactions), AOA does not partition back to the gas phase. The rates of AOA produced from LSVOCs and LSVOCs from HSVOCs are defined as:

$$\frac{dAOA}{dt} = \sum_{j=1} f_j \left[\sum_{s=1}^{S} LSVOCj \times k_{i,s} \times P_s \right] \qquad (5\text{-}9)$$

$$\frac{dLSVOC_J}{dt} = HSVOCj \times k_{i,s} \times P_s \qquad (5\text{-}10)$$

where f_j is the conversion factor from ppmV to μgm^{-3}. By adding aged aerosol, we add one more volatile bin (non-volatile organic aerosol) to low and high volatile organic carbon, which is similar to but a simpler version of the non-volatile organic compounds as explained in Donahue et al. (2009).

5.2.4 Case Studies

The modified SOA module was tested using data obtained during July 2001 and January 2002 for the continental U.S. Three test cases are performed: a base case using default CMAQ version 4.5 ($CMAQ_{base}$), a simulation using the modified SOA module with ΔH = 40kJ/mol (ΔH40), and the modified module with ΔH=156kJ/mol (ΔH156). ΔH156 was performed to test the impact of ΔH on aerosol yields. The modeling domain covered the continental U.S. and portions of Mexico and Canada with a 36-km horizontal resolution (Figure 5.1). Inputs were developed using the NCAR's fifth generation Mesoscale Model (MM5), version 3.5.3 (PSU/NCAR, 2003), the Sparse Matrix Operator

Kernel for Emissions (SMOKE), version 2.1 (EPA, 2004) (anthropogenic emissions), BEIS version 3.12 (biogenic emissions) (CMAS, 2007).

$PM_{2.5}$ measurements from the Speciated Trends Network (STN), the Interagency Monitoring of Protected Visual Environments (IMPROVE, 1995), the Assessment of Spatial Aerosol Composition in Atlanta (ASACA) (Butler et al., 2003) and the Southeastern Aerosol Research and Characterization (SEARCH) (Edgerton et al., 2005) were used to evaluate the CMAQ simulation. CMAQ simulation performance is evaluated in six regions (Figure S1) using mean fractional biases (MFB) and errors (MFE) suggested in Boylan et al. (2006). Regional performance is analyzed due to the large spatial variation of SOA precursor emissions. For hourly measurements, continuous $PM_{2.5}$ observations at Yorkville, Georgia, SEARCH monito are used.

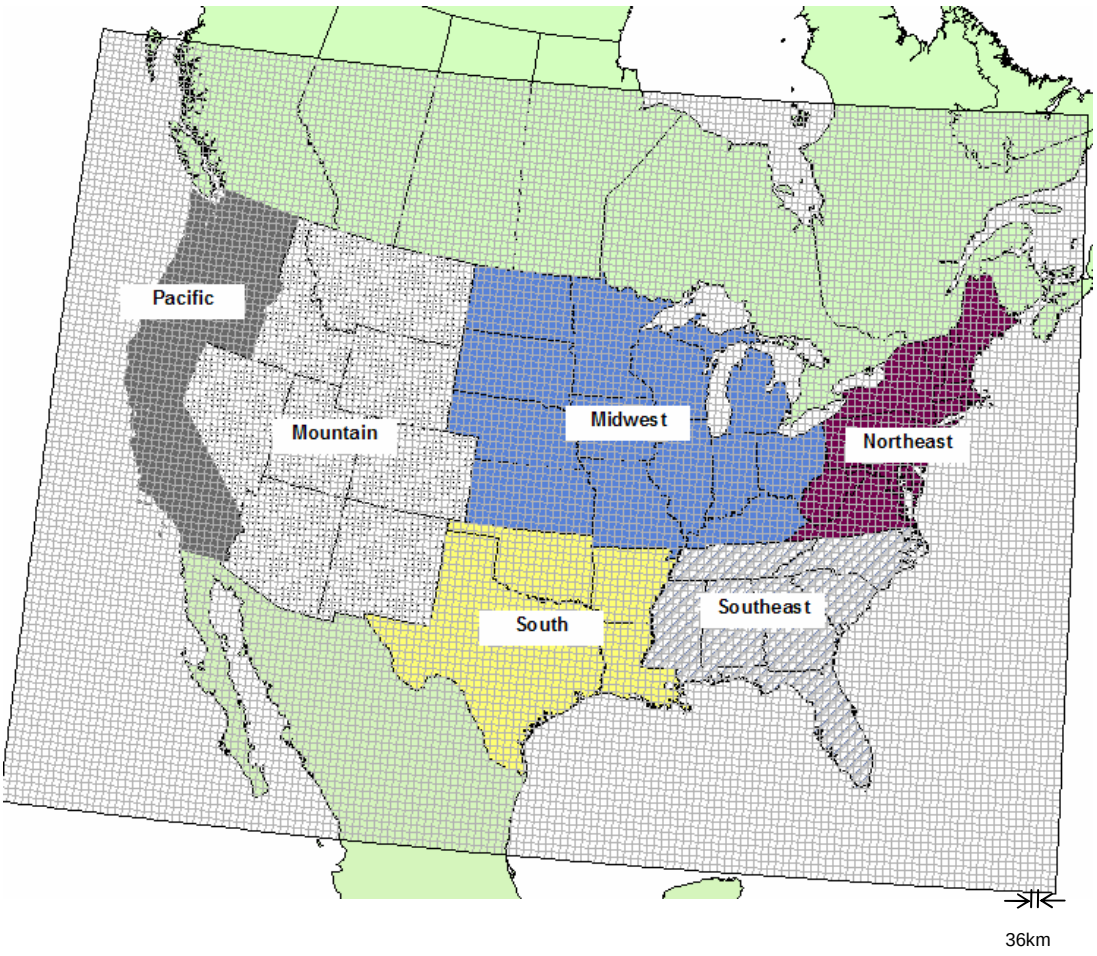


Fig 5.1 Domain of CMAQ modeling (36km resolution) and six regions in US

5.3 RESULTS AND DISCUSSION

5.3.1 Changes in α_{TRP} and $C_{sat,TRP}$ of monoterpenes

Modified α_{TRP} and $C_{sat,TRP}$ have different impacts on monoterpene SOA by region (Figure 1). $\alpha_{L,TRP}$ decreases in the Southeast and the Midwest, and rises by 30% around the Great Lakes due to increased contributions of APIN and BPIN. $\alpha_{H,TRP}$ decreases in the South, Southeast and parts of the Pacific regions and is higher in the Northwest. $C_{sat,L,TRP}$ decreases in July in half of the U.S., but increases in January around the Great Lakes, indicating that the new SOA module will lead to higher levels of monoterpene SOA in July, and less in winter in most U.S. areas. $C_{sat,H,TRP}$ increases in most U.S. areas and in both seasons, except in parts of Georgia, Florida and the Great Lakes. There are no significant spatial variations in α_{TRP} and C_{sat} in Canada and Mexico because detailed biogenic source data that matches with source categories in BEIS 3.12 are unavailable. SOA concentrations change along with α_{TRP} and $C_{sat,TRP.}$

Updating α_{TRP} and $C_{sat,,TRP}$ leads to decreases in the northeastern U.S. by 2 to 3 $\mu g/m^3$ and increases in parts of the Southeast, Mexico and Canada in summer. Decreased OC simulations in the Northeast reduce discrepancies between CMAQ simulations and observations, since increased simulated OC in $CMAQ_{base}$ is found in July in Washington and Idaho. In winter, SOA decreases in most areas with updated α_{TRP} and $C_{sat,,TRP}$.

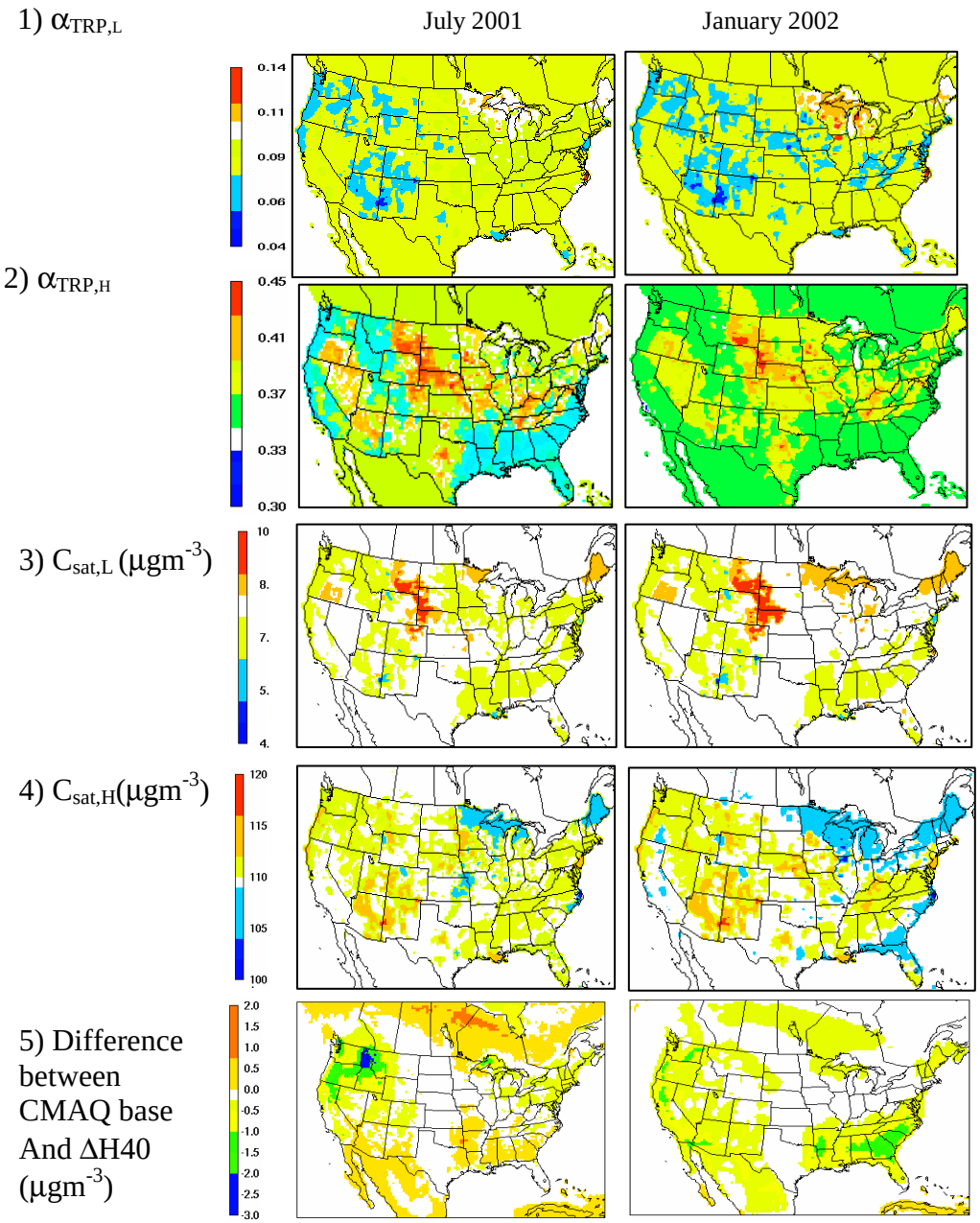


Figure 5.2 Changes in mass stoichiometric coefficients and saturation vapor pressures for low-volatility SVOC (LSVOC) and high-volatility SVOC (HSVOC) at T=298°K and M_0 = 10 μgm^{-3}. Areas in white mean little or no change in α and C_{sat} from values calculated using the default mass composition of five species in CMAQ version 4.5.

5.3.2 SOA from monoterpenes

Monoterpene concentrations (C_{TRP}) and SVOCs produced from monoterpene reactions ($SVOC_{TRP}$) in winter in the Southeast were less than one-quarter of those in the summer due to lower emissions and SOA_{TRP} fell accordingly (Figure 5.4). However, seasonality in SOA_{TRP} is less than that in C_{TRP}, partly due to a lower $C_{sat,TRP}$ in winter. SOA_{TRP} in **ΔH40** was decreased in both seasons compared to **ΔH156** as aerosol yield was halved (7% to 3%) because of the lower ΔH. Using detailed monoterpene species lowered α_{TRP}, and higher $C_{sat,TRP}$ in many areas, reducing SOA_{TRP} further. Decrease in SOA_{TRP} are largest in parts of the Pacific region. Some parts of $SVOC_{TRP}$ (less than 2% of $SVOC_{TRP}$) moved to multi-generational SVOC instead of participating in gas/particle partitioning to SOA_{TRP}, thus slightly decreasing SOA_{TRP}.

Lower ΔH significantly reduces aerosol yields in both seasons (Figure S4) to less than half of those measured in chamber studies. In summer, aerosol yields decreased from 70% (ΔH156) to 15% (ΔH40) for sesquiterpene, 11% (ΔH156) to 3% (ΔH40) for monoterpene. In winter, they decreased from 100% (ΔH156) to 50% (ΔH40) for sesquiterpene and 11% (ΔH156) to 6% (ΔH40) for monoterpene (Figure 5.6).

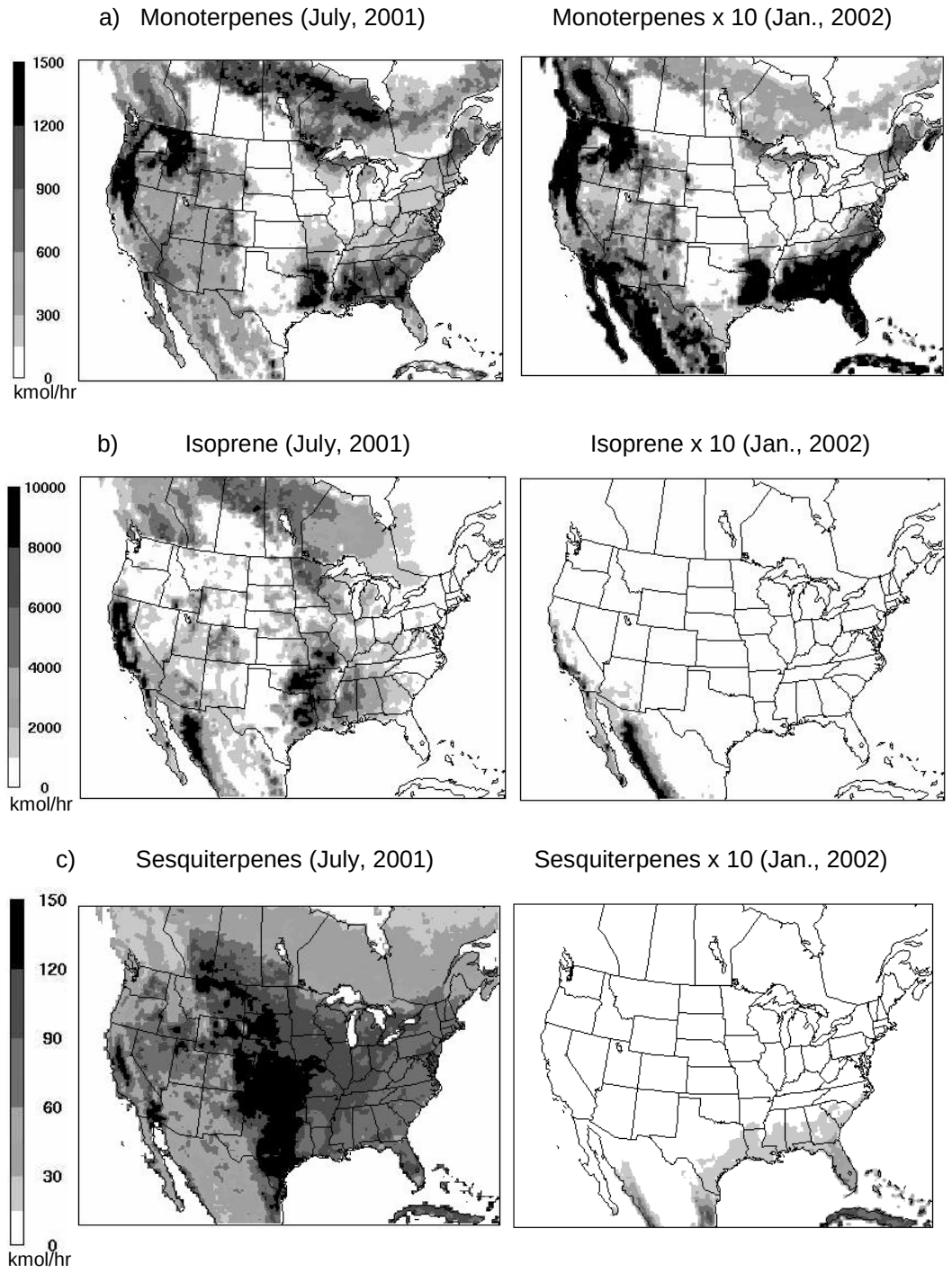


Figure 5.3 Averaged hourly (a) isoprene, (b) monoterpene, and (c) sesquiterpene emissions during July 2001 (Left) and January 2002 (Right). Note that values of January 2002 is multiplied by 10 for better display.

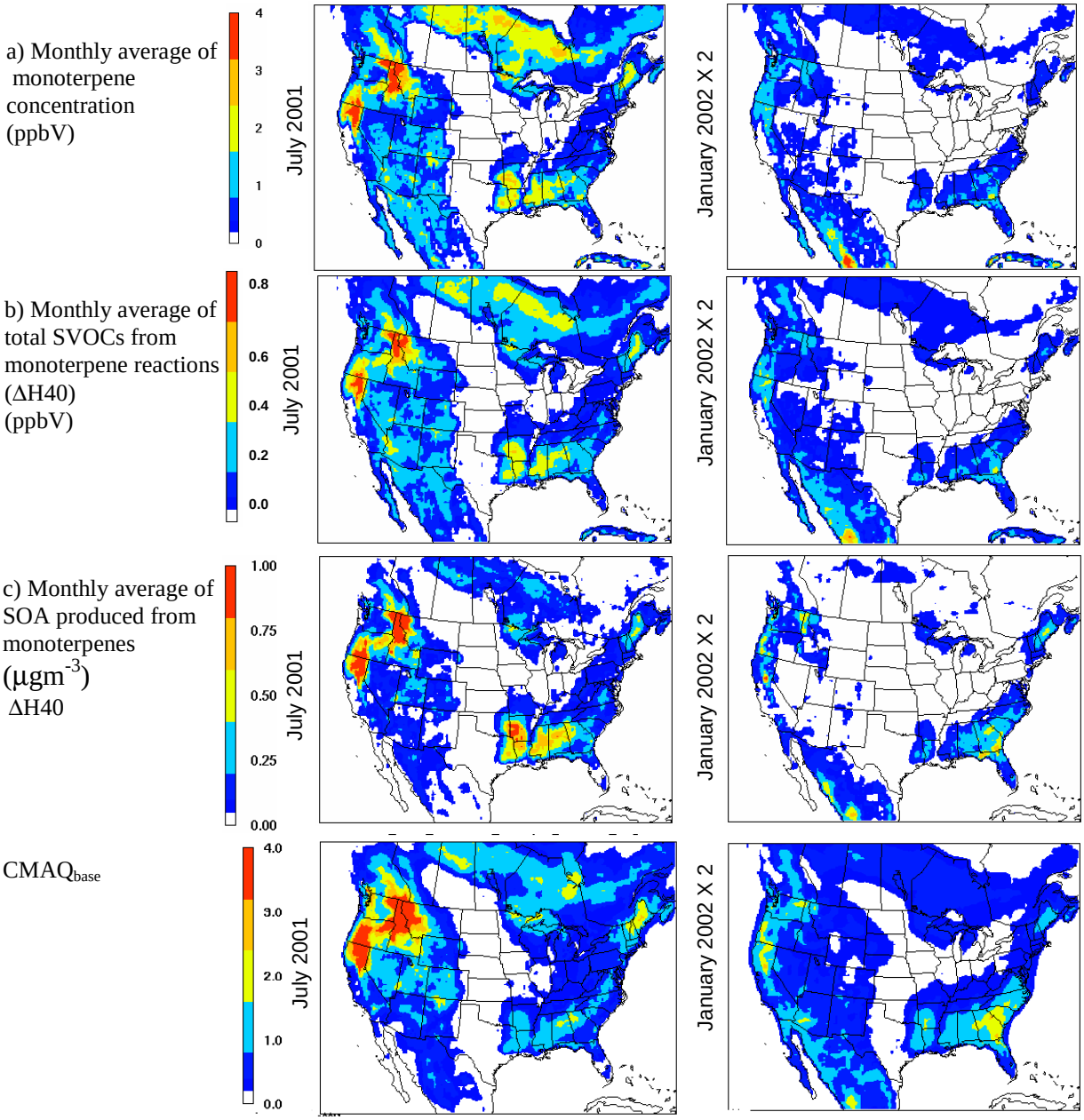


Figure 5.4 Monthly average concentrations of a) monoterpene, b) total SVOCs from monoterpene reactions, c) SOA produced from monoterpene in CMAQ base case ($CMAQ_{base}$) and the updated SOA module (ΔH40).

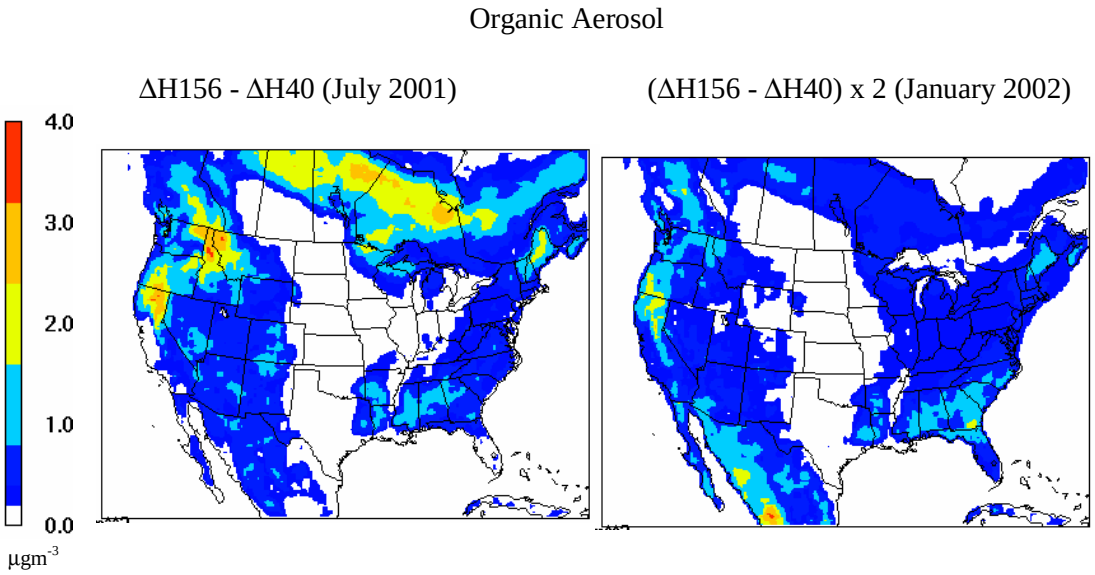


Figure 5.5 Differences in secondary organic matter between ΔH156 and ΔH40 cases. Positive values indicate decreases of SOA in the ΔH40 case.

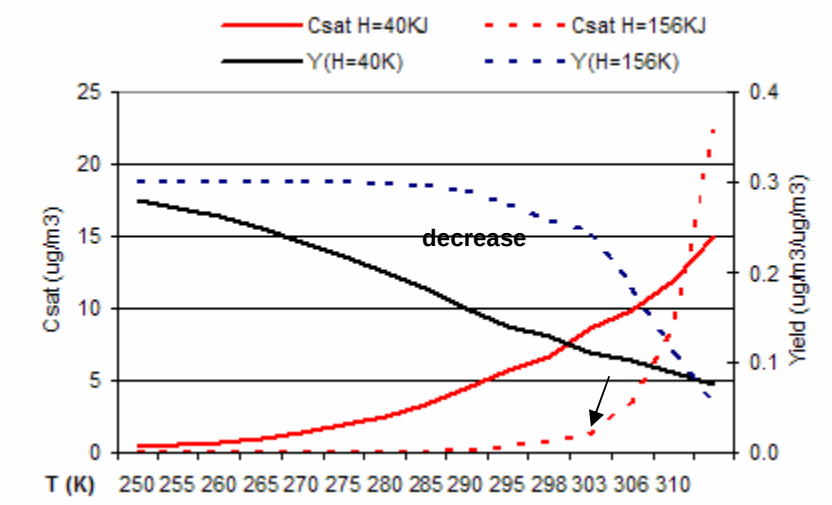


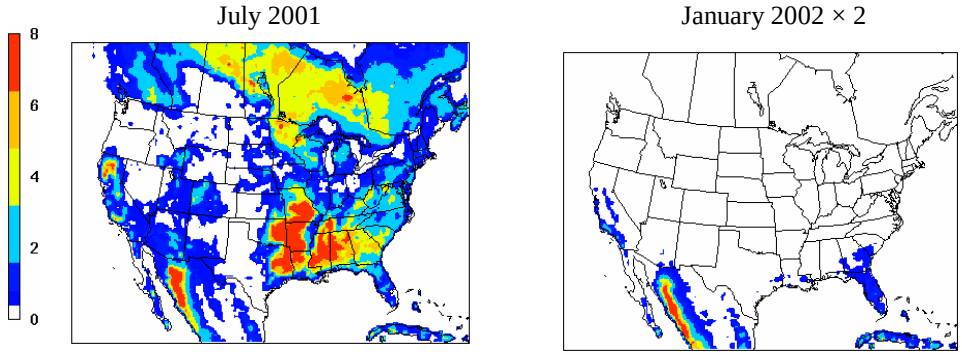
Fig. 5.6 Aerosol yield and C_{sat} changes at different temperature when enthalphy of vaporization ΔH = 156KJ/mole (dashed lines) and 40KJ/mole (solid lines). Assumptions: Reference temperature for C_{sat}, T_0 = 298K, M_0 = 5μgm^{-3} and alpha = 0.3.

5.3.3 SOA from isoprene and sesquiterpenes

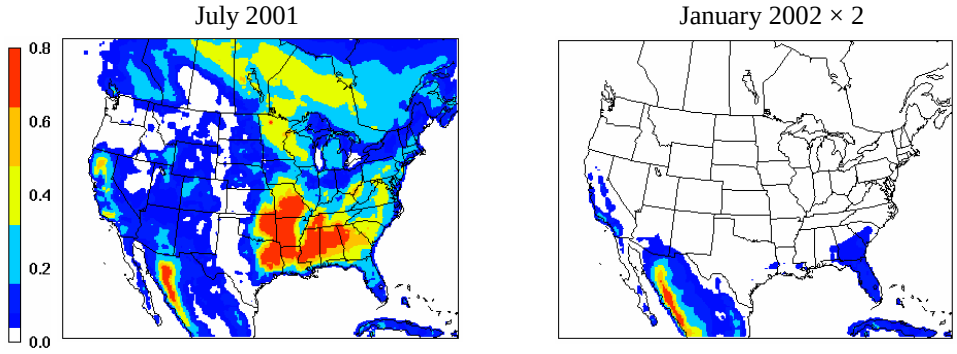
Isoprene emissions are abundant in the southeastern U.S. and low in the Midwest and West in the summer. An average concentration of isoprene in the summer in the Southeast is about 1.5 ppbv and $SVOC_{ISP}$ is about 0.4 ppbv. In winter, isoprene emissions are close to zero in the modeling domain, except for parts of Florida and Mexico. Consequently, SOA_{ISP} is minimal except in these two areas (Figure S6). Sesquiterpene concentrations are about 3% of monoterpenes in mass concentration, but the regional average of SOA_{SQT} concentration is about 10% of the SOA_{TRP} values in July 2001, because of a larger α_{SQT} (1 for $SVOC_{SQT}$ vs. 0.1 for $LSVOC_{TRP}$ and 0.3 for $HSVOC_{TRP}$) and very high reaction rates of sesquiterpenes with OH and ozone. Levels of $SVOC_{SQT}$ are similar to or larger than those of gas phase SQT as the characteristic time for sesquiterpenes in atmosphere is less than a few hours (with OH) or minutes (with ozone) (Atkinson et al., 1995). Concentrations of sesquiterpenes and $SVOC_{SQT}$ in January are about 10% of those in July due to decreased emissions.

Comparing concentrations of SVOCs with C_{sat} can explain how SOA will change with a different α or C_{sat}. SOA_{TRP} will increase with a larger α or a smaller C_{sat} as $SVOC_{TRP}$ in the atmosphere is lower than C_{sat}: approximately 0.2ppbv in summer (2 μgm^{-3} at 298K), while C_{sat} is 7 μgm^{-3} for LSVOC and 106 μgm^{-3} for HSVOC. On the other hand, $SVOC_{ISP}$ is 0.04ppbv (0.7 μgm^{-3} at 298K), which is similar to $C_{sat,ISP}$ (0.6 μgm^{-3}). Consequently, an increase of α will directly raise SOA_{ISP}, while smaller $C_{sat,isp}$ has less impact than α. $SVOC_{SQT}$ is 0.03 ppbv (0.3 μgm^{-3} at 298K) while $C_{sat,SQT}$ is 11 μgm^{-3}. A lower $C_{sat,SQT}$ will increase $SOA_{SQT,}$ while α_{SQT} is already the maximum value of α, which is 1.0.

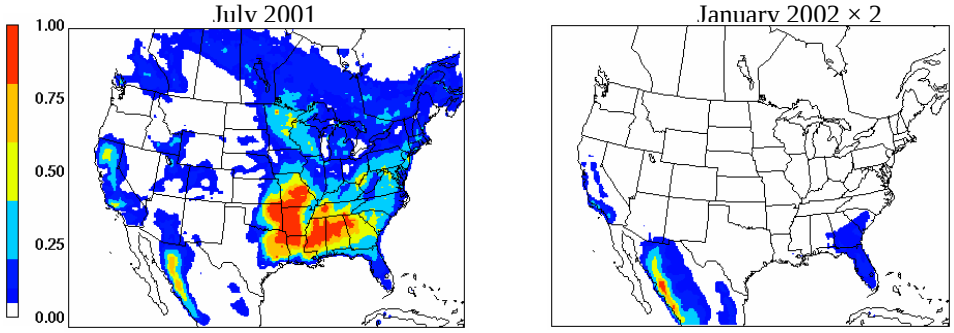
a) Monthly average of isoprene concentration (ppbV)



b) Averaged total SVOCs from isoprene reactions (ppbV)



c) Averaged SOA produced from isoprene (μgm^{-3})



μgm^{-3}

Figure 5.7 Monthly average concentrations of a) isoprene, b) total SVOCs from isoprene reactions, c) SOA produced from isoprene.

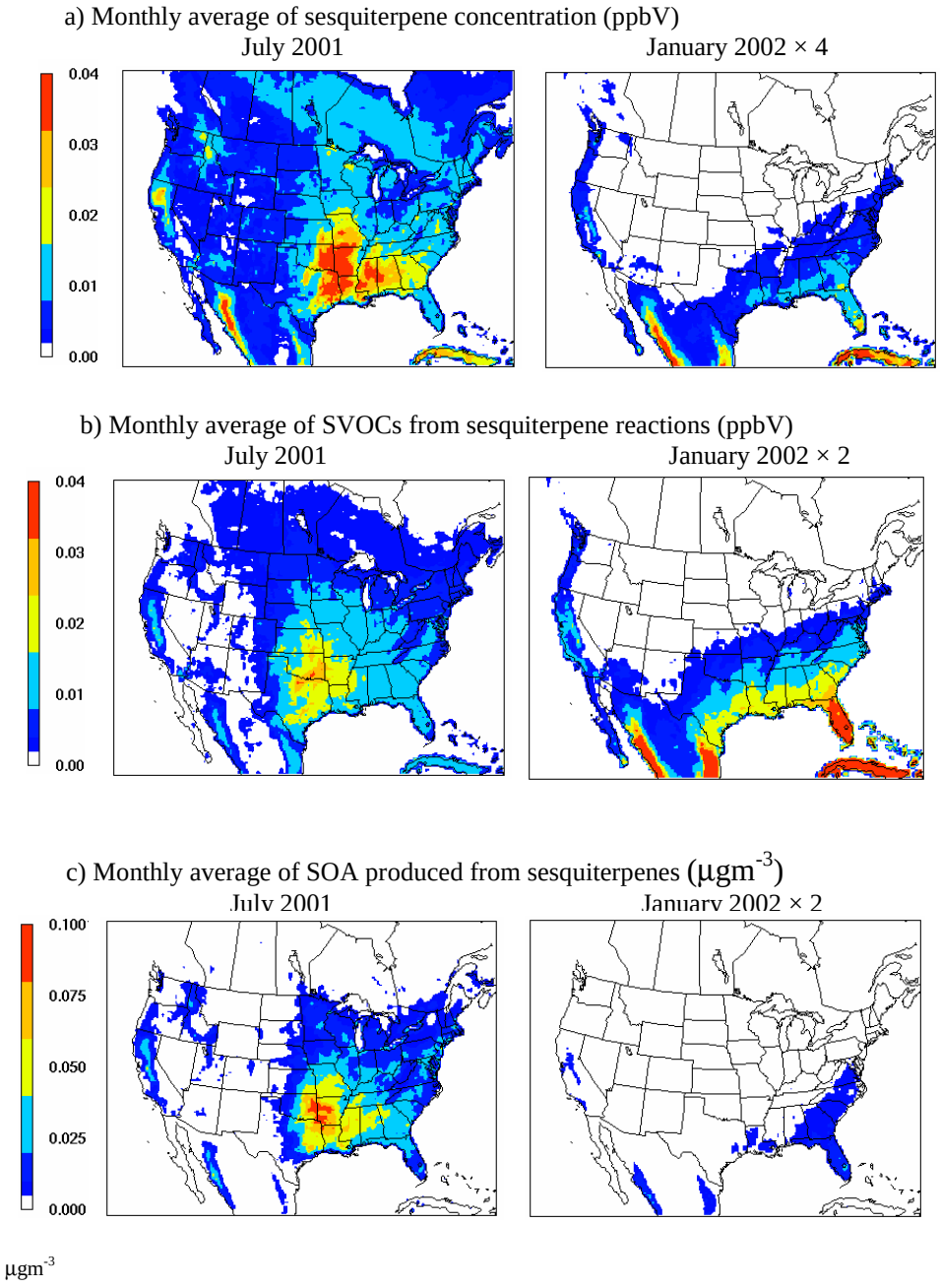


Figure 5.8 Monthly average concentrations of a) sesquiterpenes, b) total SVOCs from sesquiterpenes reactions, c) SOA produced from sesquiterpenes.

5.3.4 Aged aerosol

Aged aerosol is simulated to be a major component of SOA in both seasons. The spatial distribution of AOA is quite different from other first generation SOA (FSOA), which are regionally distributed according to their emission sources. AOA is more widely spread throughout the continental U.S., as it does not evaporate back to a gas phase SVOC. AOA originating from biogenic sources is dominant in both periods; 45% from monoterpenes, 30% from isoprene, 23% from sesquiterpenes and 2% from anthropogenic precursors in July. Anthropogenic SOA contributions increase in winter to 7% of total AOA due to less emissions of biogenic precursors.

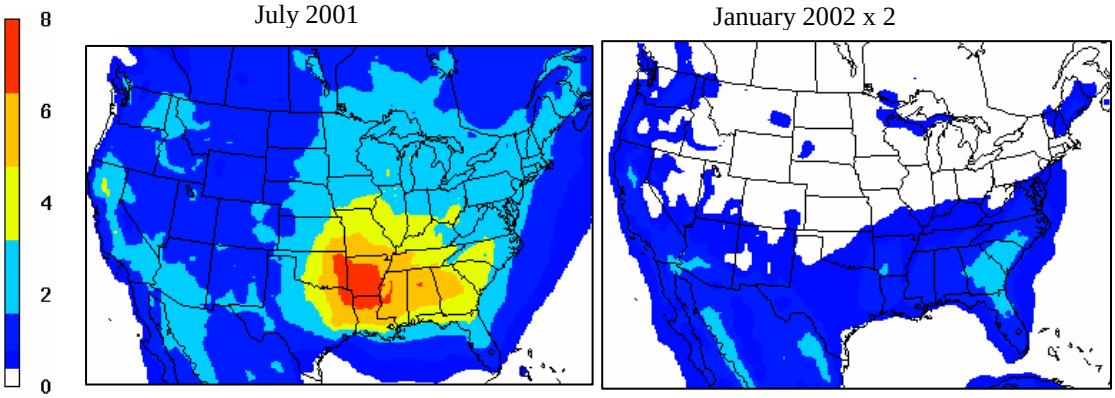


Figure 5.9 Monthly average concentrations of aged aerosol (μgm^{-3}). Note that the scale of January is doubled for presentation.

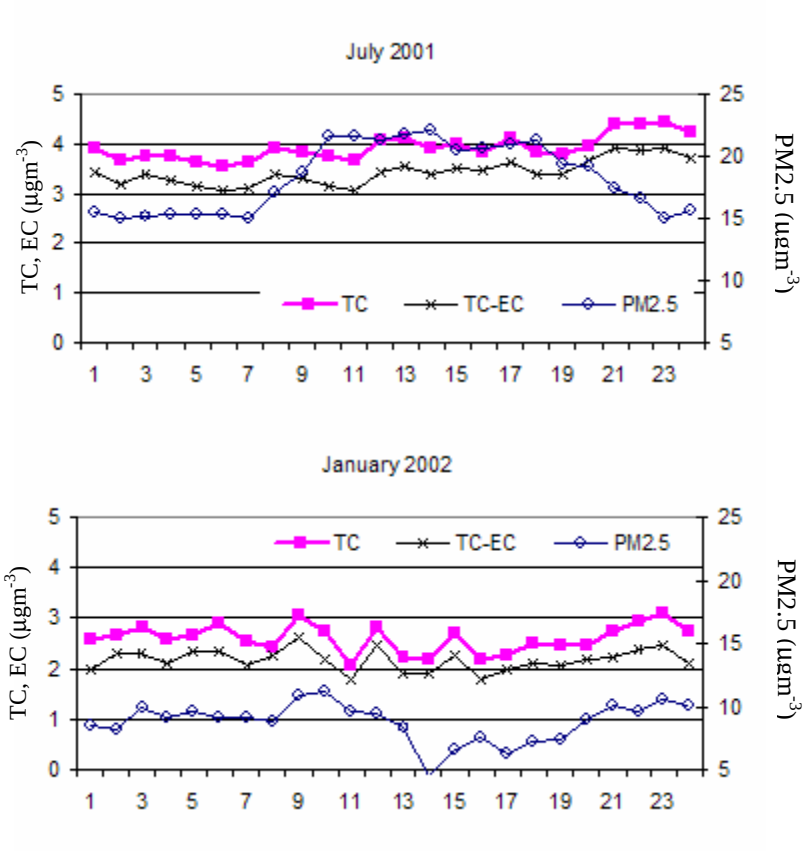


Figure 5.10 Monthly average of hourly measurements of $PM_{2.5}$ (on right Y-axis), total carbon (TC) and total carbon minus black carbon (TC-EC), at Yorkville, Georgia.

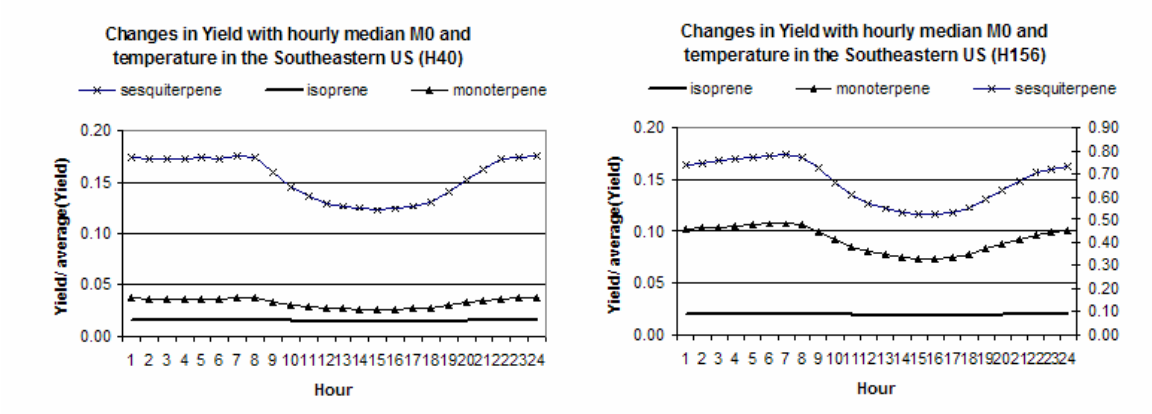


Figure 5.11 Diurnal changes in aerosol yields of monoterpenes, sesquiterpenes and isoprene in the Southeastern U.S.

5.3.5 Changes in performance

OC simulation performance significantly improved using the extended SOA approaches, especially in the South, Southeast and Pacific areas, where biogenic emissions are abundant (Figure 5.12). Negative MFBs in the eastern U.S. become close to or larger than zero, and NMEs decreased. A modest high bias from biogenic SOA appears to be the main reason for biases in the Mountain region. MFBs of organic carbon in Arizona and Utah in July are largest. 22 out of 25 stations in those states are located in National Parks, with average OC observations at 14 stations below $1\mu gm^{-3}$, while average OC simulations range from 1.5 μgm^{-3} to 4 μgm^{-3} at the same locations. In January, the updated CMAQ showed little improvement.

Simulating multigenerational aerosol is associated with very uncertain parameters. The chemical kinetics used to estimate multigenerational SVOC are not from measurements and the products and their associated properties are not known. MFB without multigenerational aerosol is -60% in summer and -40% in winter domain wide. MFBs with multigenerational aerosol are 20% (summer) and -20% (winter). MFBs for $CMAQ_{base}$ are -30% in summer and 2% in winter, thus the updated SOA module without multigenerational aerosol leads to decreases in overall SOA formation. Reducing ΔH from 156kJ/mol to 40kJ/mol lowered SOA in most regions.

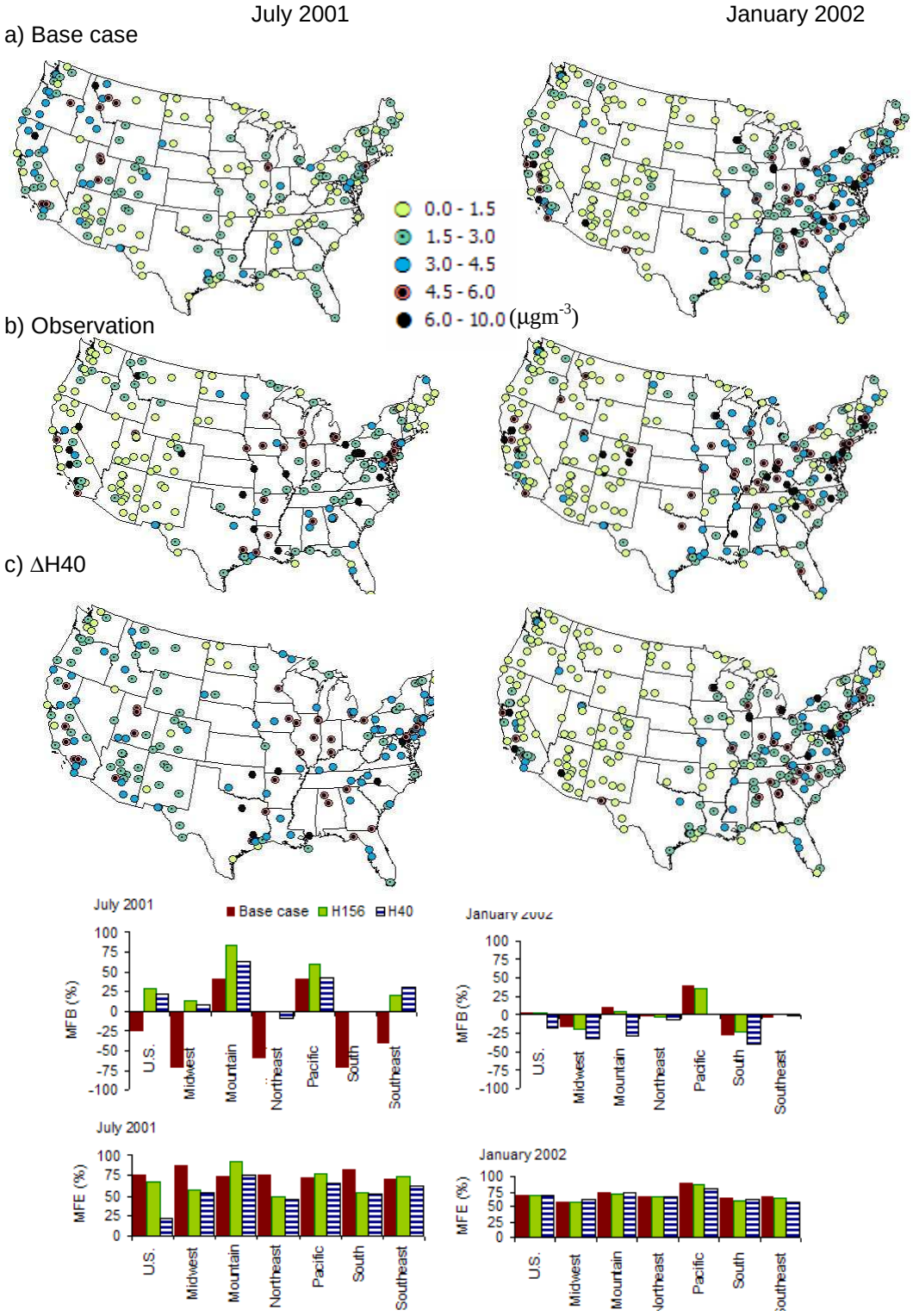


Figure 5.12 Simulated and measured organic carbon in July 2001 and January 2002. a) CMAQ base case, b) measurements and c) CMAQ improved (ΔH=40kJ/mol). Bar charts show NMFBs in each region.

5.3.6 Diurnal changes

Hourly OC measurements in 2004 at Yorkville (YRK, located in a rural area) in Georgia have little diurnal variation in July and January (Figure 5.10), while simulated SOA in prior CTM simulations found quite different diurnal patterns, especially in summer. Simulated SOA concentrations decrease about 60% during the day, leading to an overall decrease in organic aerosol (OA). However, observed OA increases. The decrease in OA simulations is driven by increasing planetary boundary layer (PBL_ height, the lower K_{om} due to higher temperature during daytime (20% decrease) and diurnal variations in primary organic carbon (10% decrease). Monoterpene, HUMUL and OSQT have the same trend, with a minimum around 14:00, but isoprene and CARYL show less diurnal variation mainly because of large increases in emissions. In winter, primary organic carbon is the dominant source of ambient OC, and both simulated and measured OC represent decreases in day time concentrations.

Aerosol yield depends on the concentration of organic carbon pre-existing in the air, and the sum of primary organic and N-generational aerosol is used as M_0 in this study. Decreasing ΔH from 156kJ/mol to 40kJ/mol lowered SOA concentrations during the daytime (Figure 5.11). Aerosol yields of sesquiterpenes decreased the most, with the difference between the maximum and minimum aerosol yield [Y_{diff}] changing from 0.25 (156kJ/mol) to 0.05 (40kJ/mol), followed by monoterpenes (0.03 to 0.01), and with no significant changes in isoprene aerosol yields.

Discrepancies between simulated and observed concentrations suggest that we are missing some precursors that are highly reactive during the daytime, or some factors that may lower the C_{sat} during daytime, or that emissions (like isoprene) are actually increased during daytime. Adding multigenerational aerosol and SOA_{ISP} reduces diurnal variations because multigenerational aerosol is very stable in the air (i.e., does not evaporate back to gas phase) and isoprene emissions increase during the daytime. N-generational aerosol is the only aerosol that shows a diurnal pattern similar to the measurements, suggesting that SOA precursors or SOA products that have similar properties as N-generational aerosol are missing in the current CTM. Biogenic emissions or aerosol yield during daytime may be higher than current estimates, or that the SOA products from further reaction are very stable (due to heavy molecular weight or low saturation vapor pressure) and do not

evaporate back in the daytime. An increase in primary OC in daytime may raise aerosol yield, but this is unlikely to be responsible because there is no evident source that emits OC significantly more during the daytime.

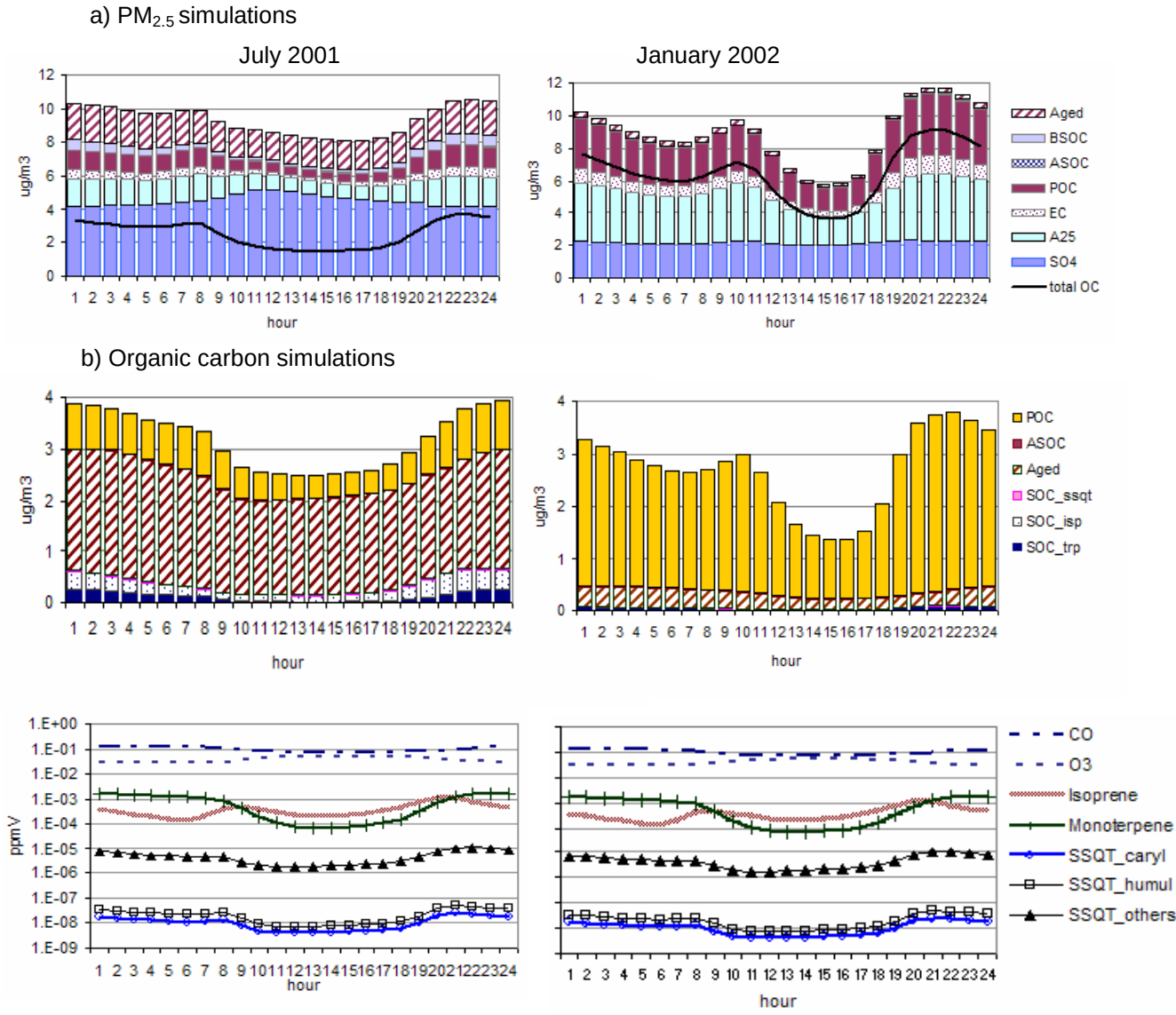


Figure 5.13 Monthly median of hourly simulations in the southeastern U.S. of a) $PM_{2.5}$ (EC: elemental carbon, NOC: N-generation aerosol, BSOC: biogenic secondary organic carbon, POC: primary OC, ASOC: anthropogenic SOC, SO4: sulfate and total OC), b) detailed OC (SOC_SSQT: BSOC from sesquiterpene, SOC_isp: BSOC from isoprene, SOC_trp: BSOC from monoterpene, and c) gas phase species (SSQT_caryl: β-caryphyllene, SSQT_humul: α-humulene, SSQT_others: other sesquiterpenes)

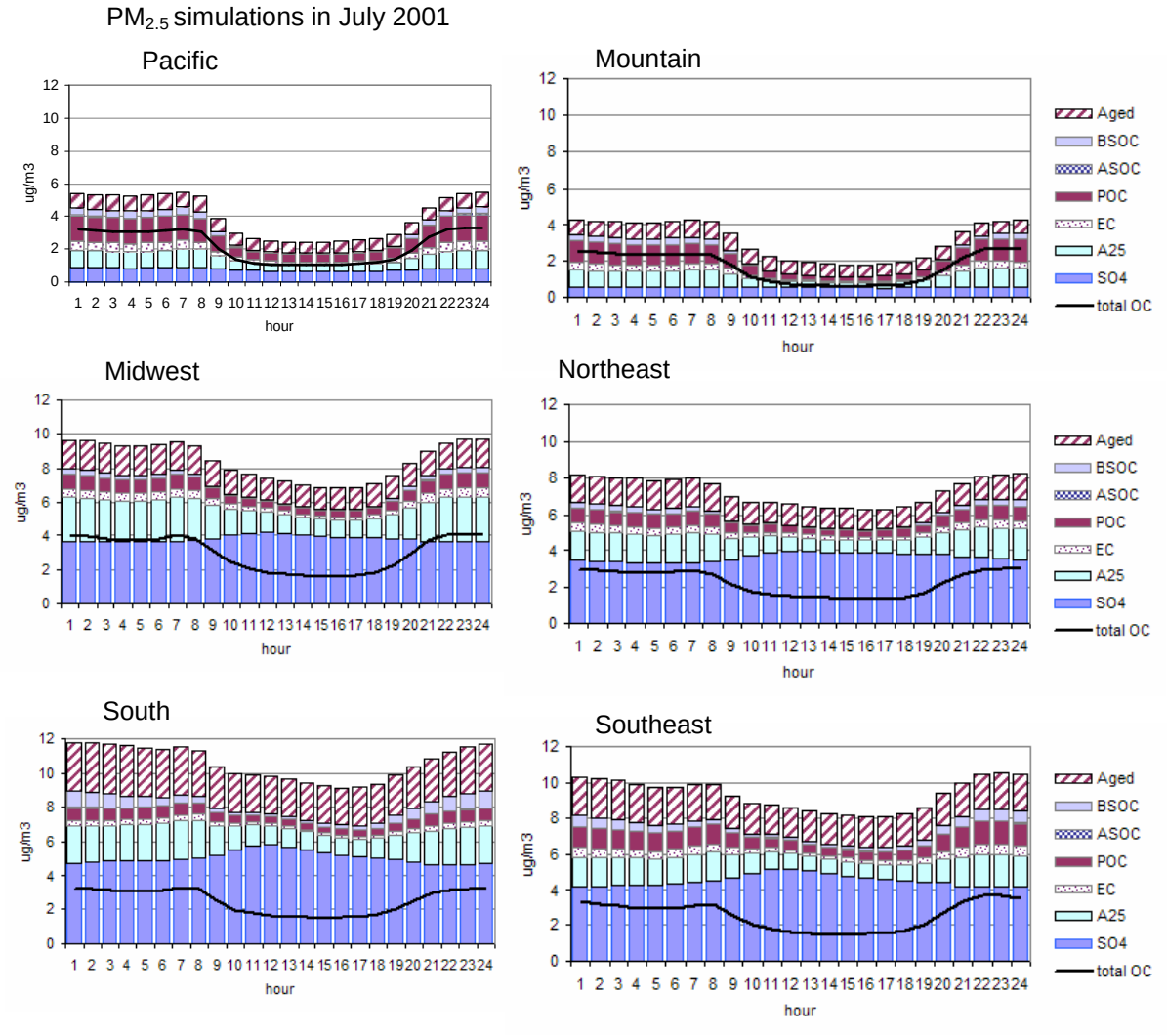


Figure 5.14 Monthly median of hourly $PM_{2.5}$ simulations in six regions. (EC: elemental carbon, NOC: N-generation aerosol, BSOC: biogenic secondary organic carbon, POC: primary OC, ASOC: anthropogenic SOC, SO4: sulfate and total OC) in July 2001

5.4 CONCLUSION

A more detailed approach to simulate secondary organic aerosol formation from monoterpenes, isoprene, and sesquiterpene reactions is developed, along with capturing more stable SOA produced from additional oxidation of semi-volatile organics. Using detailed monoterpene species emissions to estimate lumped mass stoichiometric coefficients and saturation vapor pressures tends to decrease SOA, while additional aerosol from isoprene and sesquiterpenes doubled the amount of biogenic SOA in the south and southeast regions of the U.S. Aged aerosol from stable SOA tripled SOA amounts in most regions. Overall, the modified SOA mechanism significantly improved OC simulation performance in both the summer and winter. SOA contributes 70% of total organic carbon in the Southeast with modified SOA module, which is similar contribution found by Weber et al. (2007) in Georgia and Lewis et al. (2004) in summertime using modern carbon analysis. In particular, the modified SOA module finds less diurnal variation than the original CMAQ module, especially in the summer, mainly due to aged and isoprene organic aerosol. Only aged and isoprene aerosol show a similar diurnal pattern to the OC measurements (i.e., no significant drop during daytime) in the summer. This result suggests 1) the CMAQ version 4.5 is missing precursors or mechanisms that are more stable than the current species, and/or 2) precursor emissions are underestimated during the daytime. C_{sat} of semi-volatile organic carbon decreases in daytime lower than those are measured in chamber studies, considering that SOA is oxidized in a longer period of time than experiments. Lowering the enthalpy of vaporization (from 156kJ/mol to 40kJ/mol) increased C_{sat}, but aerosol yield does not

change as much, since it is limited by the mass stoichiometric coefficient and is proportional to the inverse of C_{sat}. A smaller enthalpy of vaporization decreases overall SOA concentration, but it has little impact on seasonal or diurnal variations. This supports the importance of the consideration of aged and isoprene-derived organic aerosol formation to improve hourly OC simulations in CMAQ.

(This Page Intentionally Left Blank)

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 MAJOR FINDINGS

This thesis presents applications of detailed source apportionments of fine particulate matter to conduct chemical transport model and employed the results both for better identification of important sources of PM2.5 and for their impacts on improving performance of air quality modeling.

6.1.1 Source apportionment of fine particulate matter

PM2.5 contributions from twenty-eight sources that cover 97% of primary PM2.5 emissions were separately simulated in the Community Multiscale Air Quality (CMAQ) model to quantify the impact of each source. Apportioned source contributions, along with ionic PM2.5 species, were compared with PM2.5 measurements as well as with source impact estimates derived using receptor-oriented models. Evaluating source apportionment results of PM2.5 is not easy since there is no direct measurement of source contribution in the air. However, some consistent discrepancies between simulations, measurements and receptor-model results have clearly shown biases and errors in air quality modeling. First, weak correlation coefficients between wild fires and observed primary PM2.5 in July 2001 at all SEARCH monitoring sites suggest that using the annually averaged temporal profile for wildfires is insufficient to capture the changes in wildfire impacts. In January, daily information on wild fires was used, which led to a better correlation between simulations and measurements. Low (and sometimes negative)

correlation coefficients between soil dust impacts simulated in CMAQ and estimated by receptor oriented models strongly suggest the need for improvement of the fugitive dust emission estimates method in the emissions inventories. Lastly, failure to capture daily variations in PM2.5 observations in specific monitoring sites, such as Centreville and Outlying Land Field #8 demonstrates that important PM2.5 sources are poorly estimated in the emission inventories.

Sulfate simulation results in CMAQ are considered to be more accurate as a result of better emission estimates for SO_2 than for other air pollutants and the chemistry of sulfate formation is well known. However, excellent performance was not shown in the western U.S., possibly due to underestimated SO_2 emissions. These emissions may be underestimated for stationary fossil fuel combustion which is typically major contributors to SO_2 emissions or it is possible that other sources of SO_2 in this region may have been missed. CMB captures more of the temporal variation in source impacts at a specific receptor site but is likely less spatially representative than the emission-based model and vice versa. Detailed source contributions in CMAQ show more daily variations than previous studies that were conducted in the same periods and regions (Marmur et al., 2006) and strengthen usefulness of the CMAQ tracer method for health studies.

Completeness of profiles for all source types (i.e., no missing sources of air pollutants) is an important assumption in receptor models to obtain correct answers, but it is difficult to satisfy this requirement due to the high cost of characterizing the composition of emissions form the range of sources presents, as well as co-linearity in the emission composition among similar sources. Missing sources in receptor models may become a more serious problem in PM non-attainment areas because non-attainment

areas are typically under the influence of various sources of PM2.5. Meat cooking and industrial processes are important PM2.5 sources that are typically missing in receptor models not using organic molecular markers. Source apportionment results from CMAQ, which covers most PM2.5 sources, can be used to identify potentially untreated sources in receptor models at specific locations. Further, additional studies may follow as needed, such as adding a corresponding source profile or accounting uncertainties introduced by lack of sources.

6.1.2 Comparison of source contributions of PM2.5 resolved by receptor models and the chemical transport modeling

Source apportionment of PM2.5 results calculated using four receptor oriented models and CMAQ mostly have good correlation, with correlation coefficients (R) larger than 0.5. However, comparison between CMAQ and four receptor models shows that using different tracer species leads to rather different source contributions, even for the same sources. Wood combustion showed the largest discrepancy between CMAQ, the extended and regular CMB models, and the CMB model with the organic molecular markers. Results from CMB-RG and CMB-LGO agree well with each other since both models use trace metal species, but the results from CMB-MM does not match with CMAQ, CMB-RG or CMB-LGO. Differences between measured levoglucosan and potassium resulted in poor correlations among air quality models. Separation of gasoline engine exhaust and diesel engine exhaust may worsen the source apportionment in CMB, mainly due to co-linearity between two profiles.

6.1.3 Assessment of PM2.5 emission inventories

Observed trace metal concentrations at the SEARCH sites are compared to those simulated using CMAQ with a tracer technique to track primary PM2.5 sources. Regression analysis combining CMAQ simulation and observations suggests that major biases exist in emission estimates of a number of sources. The biases appear to be season and location (urban versus rural) dependent, and result not only from errors in the quantity of emissions, but from factors such as spatial allocation and temporal profiles as well. Differences between observations and simulated concentrations are such that wood burning emission estimates for July in rural areas are low by a factor of three, while they are low by a factor of two in urban and suburban areas, indicating that different sub-categories of wood combustion (i.e. fireplace and woodstove vs. residential yard waste burning) have different levels of biases. Current emission estimates of mobile source exhaust appear low by a factor of one to four, while soil dust emission estimates appear biased high by a factor of two to 30. Such findings suggest that emission activities and/or emission factors need to be modified.

In Chapter 4, simulated organic molecular markers in CMAQ are compared to observational data for July 2001 and January 2002 from the SEARCH study. Ridge regression analysis is used to estimate potential biases in emission inventories, in terms of scaling factors. Results from Atlanta, Georgia, suggest that emission estimates from mobile sources need to be increased in both summer and winter, but the amount of estimated bias varies with the source profile used, mainly due to the elemental carbon-to-PM2.5 ratio. Comparison suggests emissions from fireplaces and woodstoves need to be decreased by 50% to 80% in winter. However, measured levoglucosan, which originates

from wood combustion, is in good agreement with current simulated values, suggesting that there are significant errors in the wood burning source profile, and that the current profile may need to be updated for the Southeast.

6.1.4 Multigenerational secondary organic aerosol

Organic aerosol is typically one of the largest components of fine particulate matter (NARSTO, 2004). Chemical transport models often simulate organic aerosol lower than observed levels, possibly due to underestimated secondary organic aerosol (SOA). Here, an SOA module in CMAQ version 4.5 is updated by 1) using detailed monoterpene species emissions and 2) adding additional SOA from isoprene, sesquiterpene and photochemical oxidation of semi volatile organic carbon (SVOC). Accounting for detailed monoterpene species modifies mass stochiometric coefficients and the saturation vapor pressure of monoterpene SOA. Isoprene and sesquiterpene emissions are abundant in the southeastern U.S. and low in the Midwest and the West, resulting in the regionally limited increase in SOA. Oxidation of SVOC produces multigenerational SVOCs that partition into particle phase and contribute more than half of SOA.

The updated SOA module significantly improved CMAQ performance, especially in the summer (July 2001), with regional variations. In January, the updated CMAQ showed little or no improvement. Adding multigenerational SVOCs and isoprene SOA improved diurnal variations of simulated organic carbon, which are still larger than the measured values. SOA formation is sensitive to enthalpy of vaporization. Lower ΔH (40 kJ/mol) reduces SOA concentration than ΔH = 156kJ/mol with an aerosol yield falling by 70% in July. However, ΔH = 40 kJ/mol case enlarges seasonal and diurnal variations.

6.2 RECOMMENDATIONS FOR FUTURE RESEARCH

6.2.1 Simulating trace metals or HAPs using a CMAQ-TR method

In this thesis, only limited numbers of source profiles of tracer species are utilized in the chemical mass balance models, since comparison of CMAQ simulations with receptor models is a main focus of this research. We can extend usefulness of CMAQ-TR method by using more source profiles of trace metals and HAPs in EPA's SPECIATE program. Incorporating CMAQ source apportionment and comprehensive source profiles in SPECIATE will enable simulation of trace metals and HAPs, as well as apportionment of source impacts of those species. This, in turn, will benefit health studies as well as regulation of HAPs.

6.2.2 Further investigation of biases in emission estimates

Possible biases in emissions estimates due to biased source activities, spatial surrogates, and/or temporal allocation are identified for mobile sources, fugitive dust, and biomass burning. Alternative estimation methods for each of these sources need to be investigated in order to improve accuracy of CMAQ simulations in extended periods and regions. For example, Woodstoves may be spatially allocated not by populations, but could be allocated by wood consumption for heating. Construction dust needs to be incorporated with changes in actual construction activities using relevant statistical data, and the transportation fraction for fugitive dust needs to be calculated based on meteorological conditions, surface types, and road use.

6.2.3 Application to extended periods with higher resolution

CMAQ simulations in this dissertation are focused on one month in summer and one month in winter. Extending these two-month cases to one or more 12-month periods will provide a better idea of how PM2.5 source impacts change through the year. To capture industrial processes or mobile source contributions better, a higher resolution modeling will be insightful.

6.2.4 Levoglucosan and potassium

Both levoglucosan and potassium are tracers for wood combustion in receptor-oriented models. Biomass burning contributions estimated in CMB with organic molecular markers differs from that of CMB using trace metals, mainly due to differences between measured levoglucosan and potassium. Potassium is measured in both summer and winter, but levoglucosan is often under the detection limits in the summer, possibly because decay of levoglucosan in the summer. Levoglucosan and resin acids increase significantly in winter, while potassium shows weak seasonality. There may be additional potassium sources in summer, which is not included in CMAQ or the receptor models, other than wood burning are one of several possible reasons.

6.2.5 Assimilation of CMAQ and receptor model results

CMAQ and receptor models may be integrated to maximize their usage for health studies. This can be done by ensemble source apportionment results from different models via statistical methods or by assimilating results from one model according to findings in other models. One example of this is the assimilation of emissions estimates

or source impacts of fugitive dust in CMAQ with source contributions estimated in receptor models and application of this technique to the CMAQ modeling domain.

6.2.6 Further analysis of uncertainties in CMAQ simulation

In inverse modeling, uncertainties in CMAQ simulations may be important inputs for estimating scaling factors, since it may change adjustment factors for emission estimates and/or alter uncertainties of adjustment factors. Here, biases in CMAQ simulations are quantified using the differences between simulated and observed PM2.5 that resulted from errors in both meteorological fields and emission estimates. Using these values in a regression analysis that adjusts emission estimates may introduce errors in modified emission estimates, i.e., scaling factors using the combined uncertainties may over-correct emission estimates. Emission inventories are assumed the major sources of uncertainties in CMAQ simulations here, although uncertainties in meteorological field need to be quantified and separated from those in emission estimates for improvement of CMAQ performance.

APPENDIX A

IMPROVING THE POINT SOURCE EMISSIONS INVENTORY FOR GEORGIA[1]

A.1 INTRODUCTION

In the state of Georgia, point sources contribute to more than 5 % of emissions of all six pollutants (i.e., CO, SO_X, NO_X, VOCs, PM_{10} and NH_3). They made up approximately 96 % of sulfur dioxides (SO_2), 35 % of nitrogen oxides (NO_X), 17 % of fine particulate matter ($PM_{2.5}$), 14 % of ammonia (NH_3), 7 % of coarse particulate matter (PM_{10}), 7 % of volatile organic carbons (VOCs), and 4 % of carbon monoxide emissions (CO), as given in EPA's National Emissions Inventory (NEI) 99 database (EPA, 1999).

Point source emission inventories in the Atlanta ozone non-attainment area and counties located within 25 miles around the non-attainment area were updated in 2002 with surveyed data by Georgia Environmental Protection Division (Ga EPD) (Ga EPD, 2002). This newly updated survey covers point sources that emit more than 25 tons of air pollutants per year in the Atlanta ozone non-attainment area and those emit more than 100 tons in vicinity counties. Accuracy of NEI 99 database for point sources except the newly updated facilities, are still in doubt since most of the data in EPA NEI 99 point source database was projected from the 1980s and the early 1990s (Southwick, 2004). Projection is used to estimate changes in emissions based mainly upon growth. However,

[1] This is part of Fall line air quality study (FAQS) report, written by Jaemeen Baek, Alper Unal, Di Tian, and Armistead Russell.

changes such as addition of processes, modification of processes, or shutting down of plants is not taken into account. These changes may have great impact on emissions. Therefore, in order to assess the difference between the projected and surveyed emission estimates, and the importance of smaller point sources that are not covered by the Ga EPD survey, we developed point source emission inventory for all non-electric generating companies in Georgia excluding major metropolitan cities (Atlanta, Columbus, Macon and Augusta) for the year 2000. In this study, we focused on companies that emit more than 25 tons/year of any of the six pollutants.

Another objective of this study is to identify sources of uncertainty in point source emission inventory. Uncertainty can be defined as "a statistical term that is used to represent the degree of accuracy and precision of data" (McInnes, 2001) and is mainly due to uncertainty in emission factors or activity data. Here, we quantified uncertainties in point source emission inventory, especially resulted from incompleteness of emission inventories.

A.2 METHODS

We first developed a pre-survey in order to identify companies that emit more than 25 tons/year of any of the six pollutants. Those companies filtered in the pre-survey participated in the main survey. Microsoft Access was used to develop a database from the survey results. In the next step we did a quality assurance/quality check (QA/QC) to the database. In the cases where errors were identified in survey forms, we submitted complementary surveys. It should be noted that data input, QA/QC, and complementary

Surveys were iterated until we were confident about the quality of the data. Figure 1 summarizes the overall methodology followed in this study.

A.2.1 Pre-survey

In identifying major point sources, the Title V Permit List; Toxic Database; and Georgia Manufacturer's database were utilized, in addition to the prior EPA NEI96 inventory. All of the companies in Title V permit list and Toxics Database were included in the survey. Companies in Manufacturer's Database were screened by number of employees and type of industrial sector. This effort provided 1308 companies to be surveyed. Out of 1308 companies, 100 out of 300 companies that responded to our survey had emissions greater than 25 tons/year of the listed pollutants.

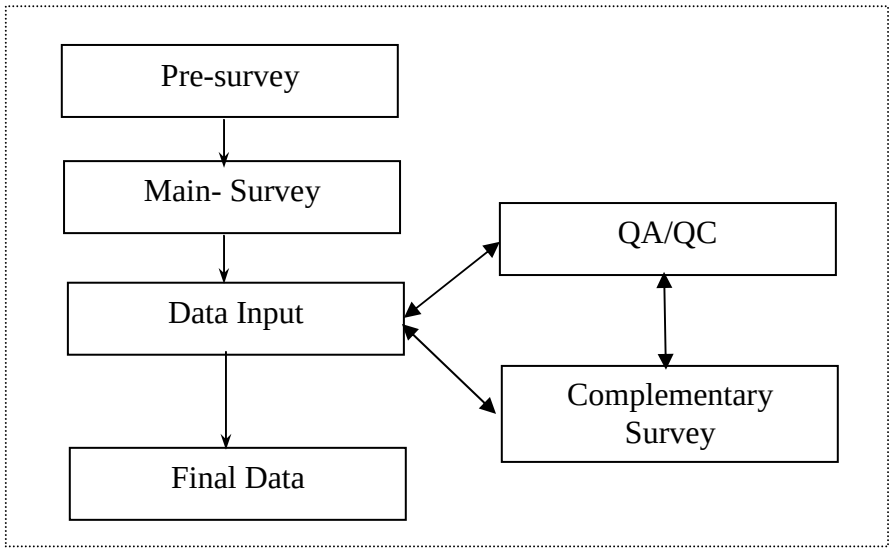


Figure A.1 Overall process of the study

A.2.2 Main Survey

Survey forms for plants with emissions larger than 25 ton/year include both plant general information and detailed information for emission estimates such as fuel burning process information, and miscellaneous process information. Forms used in this survey can be found elsewhere (FAQS, 2004).

A.2.3 Data input

Data in survey forms were entered into Microsoft Access© tables using data input program which was developed using Visual Basic for Applications (VBA). Objectives of data input program are: 1) to provide easy-to-use graphical user interface (GUI) to reduce errors during data entering, 2) to provide easy access to national emissions inventory input format (NIF) codes to reduce working time and to maintain consistency, 3) to check relations between tables (e.g., county federal information processing standards [fips], site ID, unit ID and process ID among general facility information, fuel burning process information and emission estimates information), and 4) to check range of input data based on EPA quality assurance guidelines for NIF (EPA, 1999).

a) b)

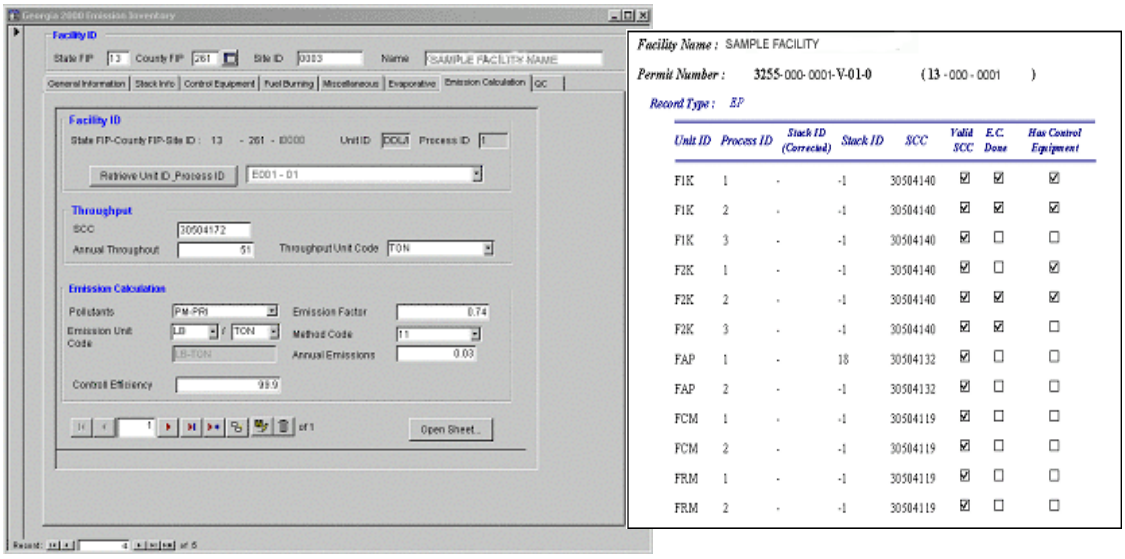


Figure A.2 (a) GUI of the data input program (b) Report generated to check completeness of surveyed data

GUI input is prepared such that it resembles survey forms, so that users can easily match the paper form with GUI (Figure A.2). The report generated by the program to check completeness of survey data (Figure A.2, b) lists all emission processes in each company, stacks and control devices connected to those processes, validity of SCC and existence of emission estimates. Additionally, users can access codes defined for NIF version 2.0, such as control device codes, emission calculation method codes, fuel type codes, pollutant codes, and unit codes. By using codes during data entering, we can maintain consistency with NIF codes and also reduce unnecessary effort for converting general notions to codes.

In order to keep all information which is needed to calculate emissions, new tables that correspond to tables in NIF version 2.0 are used for data entering and for QA/QC.

Primary key and foreign key function in Microsoft Access© are used to maintain identification of emission processes and relationships between tables. Also, to limit the range of input data, validation rules are defined as below.

A.2.4 QA/QC

QA/QC is an essential part of this study developed not only to check quality of data but also to analyze and quantify errors in surveyed data for further uncertainty analysis. Criteria for QA/QC are as follows:

- Criteria for fuel burning/miscellaneous/evaporative processes:
  - ✓ Every emission unit and process should have related stack information.
  - ✓ Every emission unit and process should have emission estimates.
  - ✓ Every emission unit and process should have suitable SCC
  - ✓ Material used for emission unit and process should match with material assigned to SCC
  - ✓ Same data reported in different tables should be same (Table 1)
- QA/QC criteria for stack information:
  - ✓ All parameter should have been reported
  - ✓ Ranges of stack parameters should match with QA guidance by EPA
- For emission calculations:
  - ✓ If companies report they use EPA emission factor method to estimate emissions, the values of emission factors should be same as those in EPA factor information retrieval (FIRE) version 6.23 (EPA, 1999)
  - ✓ Units of emission factors and units of yearly throughput should be same

- ✓ Result of emission values should be same as values calculated using reported emission factors and throughputs

Table A.1 Fields that are common in more than one table. For example, “yearly throughput” of fuel is reported both in a table about a fuel burning unit and in emission estimate table for the same unit.

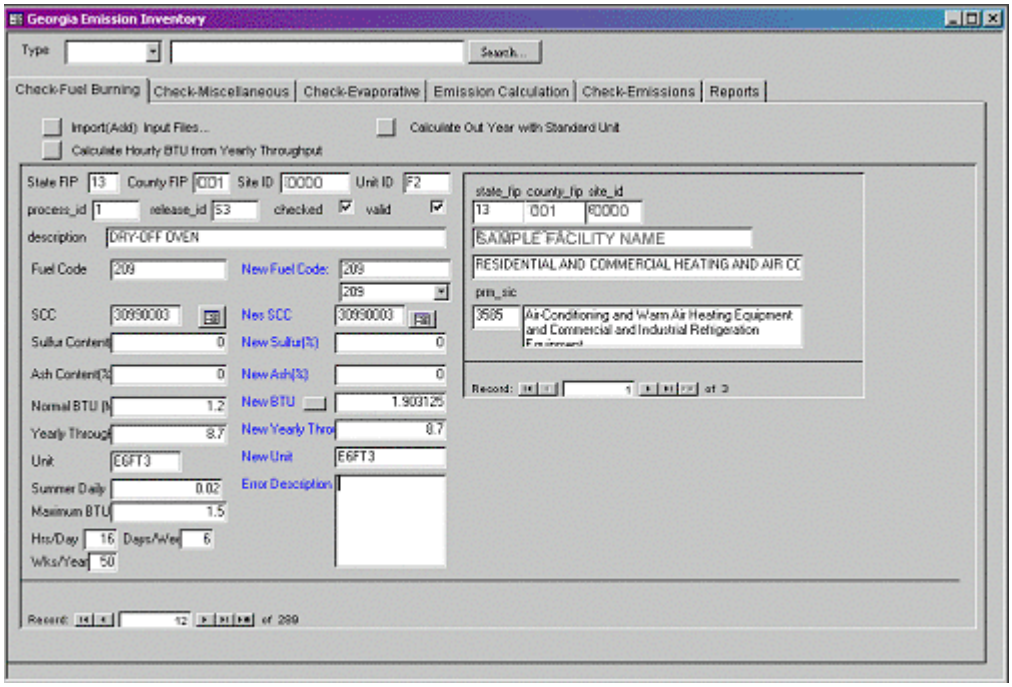
Items	Table in survey form
Yearly throughput	Fuel burning process
	Miscellaneous process
	Emission calculation
VOC emissions and uncontrolled VOC	Evaporative process
	Emission calculation
Control efficiency	Control equipment
	Emission calculation
Emission process and connected release point ID	Fuel burning process
	Stack information

Any record that fails to satisfy these criteria is corrected, and some of the corrected data are reported to companies to confirm correction. In addition to the QA/QC inherent in the data input program, a separate QA/QC program was developed. This program was developed to automatically check QA/QC criteria and to provide GUI which users can review all tables and records (Table A.2). The GUI of QA/QC program to check fuel burning processes and an example report created by QA/QC program for complementary survey is in Figure A.3.

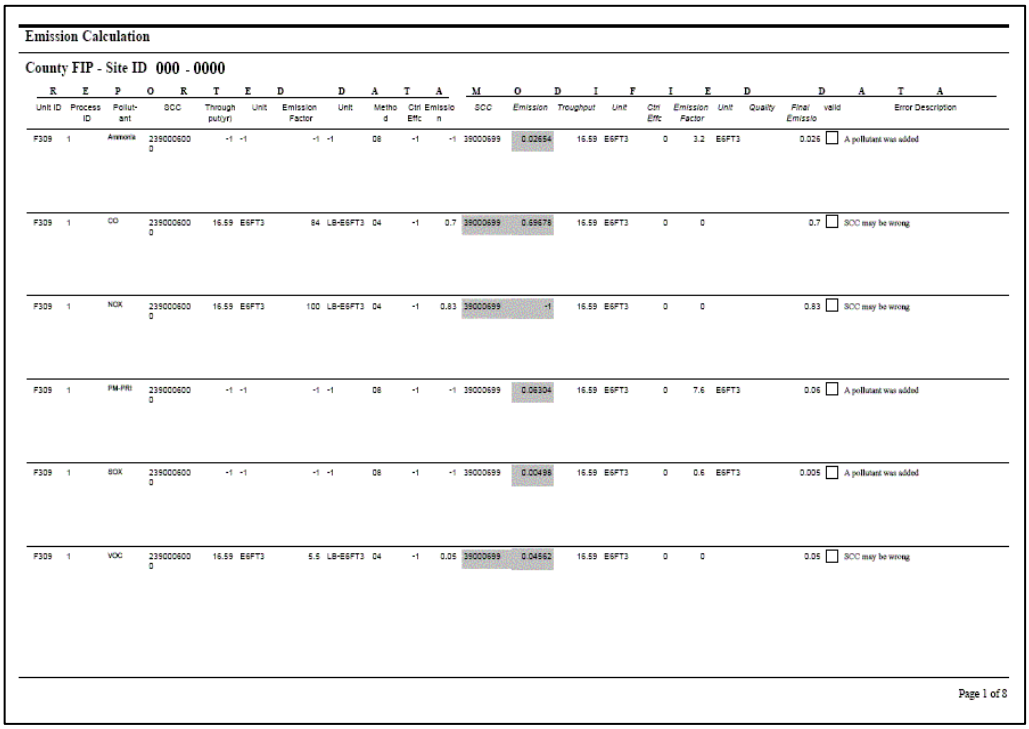
Table A.2 Components of the QA/QC program and functions

Components	Main functions
Fuel burning process	• Calculate normal BTU using fuel consumptions and operation time • Check SCC
Miscellaneous process	• Check SCC
Evaporative process	• Calculate VOC emissions based on material balance
Preparation for emission calculation	• Add emission estimates for emission processes which don't have reported emission calculation • Calculate emissions using emission factors in EPA FIRE version 6.23 • Assign control efficiencies from control equipments • Assign yearly throughputs from fuel burning processes
Emission calculation	• Check emission calculation record one by one • Check miscalculation • Assign error code and error description if needed
Reports	• Check relations between tables • Create reports for complimentary survey
Export to NIF	• Convert internal tables to NIF tables • Convert NIF tables to IDA format

a)



b)



Emission Calculation

County FIP - Site ID 000 - 0000

REPORT DATA													MODIFIED DATA									
Unit ID	Process ID	Pollutant	SCC	Throughput(yr)	Unit	Emission Factor	Unit	Method	Ctrl Effc	Emission	SCC	Emission	Throughput	Unit	Ctrl Effc	Emission Factor	Unit	Quality	Final Emissio	valid	Error Description
F309	1	Ammonia	2390006000	-1	-1	-1	-1	08	-1	-1	39000699	0.02654	16.59	E6FT3	0	3.2	E6FT3		0.026		A pollutant was added
F309	1	CO	2390006000	16.59	E6FT3	84	LB-E6FT3	04	-1	0.7	39000699	0.69678	16.59	E6FT3	0	0			0.7		SCC may be wrong
F309	1	NOX	2390006000	16.59	E6FT3	100	LB-E6FT3	04	-1	0.83	39000699	-1	16.59	E6FT3	0	0			0.83		SCC may be wrong
F309	1	PM-PRI	2390006000	-1	-1	-1	-1	08	-1	-1	39000699	0.06304	16.59	E6FT3	0	7.6	E6FT3		0.06		A pollutant was added
F309	1	SOX	2390006000	-1	-1	-1	-1	08	-1	-1	39000699	0.00498	16.59	E6FT3	0	0.6	E6FT3		0.005		A pollutant was added
F309	1	VOC	2390006000	16.59	E6FT3	5.5	LB-E6FT3	04	-1	0.05	39000699	0.04562	16.59	E6FT3	0	0			0.05		SCC may be wrong

Page 1 of 8

Figure A.3 (a) GUI of QA/QC program (b) a report for complementary survey that shows units with errors and suggested correction.

A.2.5 Complementary survey

The purpose of the complementary survey is: 1) to supplement missing information in survey form; 2) to confirm whether corrections made during QA/QC were necessary or not; and 3) to alert companies of errors in their survey forms, so that they can improve it. The complementary survey lists all relevant information with possible errors and comments on errors such that companies can re-check information and add comments to comments. Most of the complementary surveys were for the purpose of gathering more information on stacks and correcting errors in emission calculations.

A.3 RESULTS AND CONCLUSION

In this section we will summarize the results of this study and present findings from data analysis. We first present responses to our survey and then present findings from QA/QC. Later we give a quantitative summary of our study as well as comparison to the NEI 99 database.

A.3.1 Responses to the Survey

The responses to pre-, main-, and complementary surveys are listed in Table A.3. For the pre-survey, we sent forms to 1308 companies and received about 300 replies, approximately 23 percent. Out of these 300 companies, 79 of them reported that they emit more than 25 tons per year. We submitted 54 complimentary surveys which had a response percent of 55 percent.

Table A.3 Results of the paper survey and the complementary survey

	Forms Sent to companies	Replies
Pre-survey	1300	300
Survey	300	79
Complementary survey	54	30

A.3.2 QA/QC

In QA/QC we identified 12 categories for possible error for emissions estimation (Table A.4). For example, a code “00” and code “09” indicate that there is no error in emission calculation. Code “09” indicates that emission calculation comes from evaporative process information using a material balance. 71% of emission calculations were correct. The most frequent errors were the ones where there were no corresponding emission estimates for emission processes (8.6%, code 07). Next highest number of errors occurred for the cases where emission estimates for some of the pollutants were missing (7.9%, code 08).

Frequencies of errors for each pollutant were found to be different (Table A.5). In the case of $NH_{3,}$ it should be noted that most of the companies didn’t report ammonia emissions since ammonia estimates had not been included in emission inventories previously.

One interesting finding is that the third most frequent error in CO, NO_X, PM_{10} and VOC is “Using revoked emission factors”, error code 01. This error is the result of some companies using the old emission factors rather than the emission factors that were updated in 1998. Another common error was that due to reporting one combined

emission estimate for more than two emission processes (error code 11). This error contributed more than 10% for CO, NO_X and SO_X. Sometimes this type of error is inevitable because companies cannot separately measure emissions from every emission units, especially when emission units are small.

Table A.4 Error types and frequency of errors

Code		Description	Counts (%)
No Errors	00	No error	2,235 (55.0%)
	09	Emission estimate is calculated using evaporative processes information by material balance	653 (16%)
	Total		2888 (70.0%)
Errors	01	Revoked emission factors are used	236 (5.8%)
	02	Units of emission factors don't match with units of activity	21 (0.5%)
	03	Units of yearly throughput is wrong	0
	04	Typo in survey form	13 (0.3%)
	05	Miscalculation in emission estimates	18 (0.4%)
	06	Value of emission factor is wrong	12 (0.3%)
	07	Emission process doesn't have corresponding emission estimates	346 (8.5%)
	08	Emission process have corresponding emission estimates, but some pollutants are missing	324 (7.9%)
	10	Typo or omission during data entering	0
	11	Emissions from more than two emission processes are reported as one record	144 (3.6%)
	12	Wrong SCC is assigned to emission process	61 (1.5%)
	Total		1175 (30%)

Table A.5 Number of records grouped by error types

Error codes	NH_3	CO	NO_X	PM_{10}	SO_X	VOC	Total
01	-	68 (33.7%)	24 (15.4%)	71 (30.0%)	-	73 (33.3%)	236 (20.1%)
02	1 (0.4%)		5 (3.2%)	4 (1.7%)	5 (4.1%)	6 (2.7%)	21 (19.8%)
04	-	1 (0.5%)		6 (2.5%)	1 (0.8%)	5 (2.3%)	13 (1.1%)
05	-	4 (2.0%)	2 (1.3%)	7 (3.0%)	1 (0.8%)	1 (0.5%)	15 (1.3%)
06	-	1 (0.5%)	6 (3.9%)	2 (0.8%)	2 (1.6%)	1 (0.5%)	12 (1.0%)
07	41 (17.2%)	54 (26.7%)	66 (42.3%)	82 (34.6%)	42 (34.1%)	64 (29.2%)	349 (29.7%)
08	185 (77.7%)	36 (17.8%)	10 (6.4%)	32 (13.5%)	41 (33.3%)	20 (9.1%)	324 (27.6%)
11	6 (2.5%)	22 (10.9%)	27 (17.3%)	28 (11.8%)	25 (20.3%)	36 (16.4%)	144 (12.3%)
12	5 (2.1%)	16 (7.9%)	16 (10.3%)	5 (2.1%)	6 (4.9%)	13 (5.9%)	61 (5.2%)
No. of records with errors	238	202	156	237	123	219	1175
No. of total records	254	403	490	1053	411	1452	4063

After identifying the sources of errors, we corrected them through complimentary surveys as explained in the methodology section..The biggest change occurred for NH_3 and CO after corrections, with 12.7 and 12.6 percent of originally reported emission estimates respectively, whereas the smallest change occurred for SO_X, with 0.1 percent (Table A.6)

Table A.6 Changes of emissions due to corrections (ton/year)

Error Codes	NH_3	CO	NO_X	PM_{10}	SO_X	VOC
01	-	11.9	-9.7	0.4	-	-0.1
05	-	-3.7	-2.5	-0.6	-0.1	-238.3
06	-	0.0	0.0	0.0	0.3	0.0
07	4.4	317.3	879.6	67.2	5.1	525.5
08	41.0	1,661.7	190.3	7.8	9.9	9.9
11	0.7	3.3	-10.7	1.8	0.0	-72.0
12	4.3	20.7	-32.8	0.0	0.0	0.0
Grand total (ton/yr)	50.4	2,011.1	1,014.3	76.6	15.2	229.1
Changes (% of original surveyed data) after correction	12.7%	12.6%	4.4%	1.1%	0.1%	2.2%

A.3.3 Summary of Survey Results and Comparison to NEI 99

From our corrected survey results, it is found that point sources emit approximately 24,000 tons of nitrogen oxides, 20,000 tons of sulfur dioxides, 18,000 tons of carbon monoxide, 10,000 tons of VOC and 7,000 tons of particulate matter per year. Total emissions with respect to different source categories as well as percent contributions, as given with SCC level 1, are also provided. (Table A.7). External combustion boilers

contributes most to CO emissions. This is also true for SO_X. For NH_3, PM_{10}, and VOC, the biggest contribution is from industrial processes. For $NO_{X,}$ internal combustion engines are the largest contribution.

Contributions of SCC level 1 groups are compared with those in NEI 99. Table 8 shows contributions of SCC level 1 groups to emissions from non-EGU devices in NEI 99. Industrial processes are the major sources for CO, NH_3, VOC, and PM_{10}, whereas external combustion boilers are for NO_X and SO_2. The biggest difference between these values and our survey findings is in CO and NO_X. It should be noted that industrial processes are also a significant contributor for CO (i.e., 34 percent) in our survey, as seen in Table A.7. Contribution of external combustion boilers to NO_X emissions is significant in our survey as well (i.e., 33 percent). Differences between Table A.7 and A.8 may be due to dissimilarity in companies covered in both inventories and/or incompleteness of our survey.

Table A.7 Total emissions grouped by SCC level 1 (tons/year)

SCC level1	CO	NH_3	NO_X	PM_{10}	SO_X	VOC
External Combustion Boilers	10,454 (58%)	22 (5%)	7,902 (33%)	1,562 (23%)	12,115 (58%)	1,011 (10%)
Internal Combustion Engines	1,280 (7%)	9 (2%)	12,302 (51%)	119 (2%)	2,862 (14%)	1,357 (13%)
Industrial Processes	6,179 (34%)	416 (93%)	3,969 (16%)	5,111 (74%)	6,018 (29%)	7,151 (70%)
Petroleum and Solvent Evaporation	2.3 (1%)	-	3 (0%)	2 (0%)	0.0	754 (7%)
Waste Disposal	-	-	-	94 (1%)	-	-
Total	17,915	447	24,176	6,888	20,995	10,264

Table A.8 NEI 99 point source emissions from non-EGU equipments grouped by SCC level 1

SCC level 1	CO (%)	NH_3 (%)	NO_X (%)	PM_{10} (%)	$PM_{2.5}$ (%)	SO_2 (%)	VOC (%)
External Combustion Boilers	13.7	1.65	58.2	27.9	26.5	58.0	20.5
Internal Combustion Engines	0.5	0.00	12.9	0.0	0.0	0.0	0.8
Industrial Processes	85.6	98.6	28.6	72.1	74.4	42.0	43.2
Petroleum and Solvent Evaporation	0.01	0.00	0.1	0.00	0.00	0.05	35.1
Waste Disposal	0.2	0.00	0.2	0.02	0.01	0.01	0.5

In order to identify differences in the 2000 emission inventory and NEI 99, we performed a detailed comparison. For this comparison we identified common companies in both NEI 99 and our survey database. It should be noted that there may be differences between two emission inventories due to different base years, but we assumed that differences due to errors in the emission inventories are much greater than differences due to different base years. It is found that out of 79 companies in the 2000 emission inventory and 339 companies in NEI 99, 30 companies are common in both databases. Table 9 shows the changes in emissions of 30 common companies for both inventories.

The biggest difference occurred for CO, PM_{10}, NH_3, and SO_2. The change in NH_3 is very big since in NEI 99 NH_3 emissions are very small, which may be due to the fact that NH_3 emissions were not reported in earlier emissions inventories (Table A.9).

Table A.9 Change in emissions for 30 common companies. (ton/year)

	CO	NH_3	NO_X	PM_{10}	SO_2	VOC
Emissions of common companies in 2000 survey	15,304	439	18,460	3,470	18,265	6,680
Emissions of common companies in NEI 99	61,587	48	20,564	6,928	27,921	5,936
Changes (2000 survey – NEI 99)	-46,265	392	-2,082	-3,450	-9,652	809
Changes (%) to NEI 99 emissions	-75%	816%	-10%	-50%	-35%	13%

A cumulative probability plot shows the change in emissions between our survey and NEI 99 for different pollutants (Figure A.4). Each of the points in the graphs represents common companies between NEI 99 and our survey. For all of the pollutants, except NH_3 and VOC, 60 percent of the companies have differences less than zero, indicating that NEI 99 emissions for these companies are larger. Another important finding is that for NO_X, NH_3, CO, and VOC for some companies differences are more than 200 percent. In these cases, our survey emission estimate is more than two times higher than NEI 99's estimates.

Changes in emissions are mainly due to poor accuracy of NEI 99. As stated earlier, most of emission estimates in EPA NEI 99 for point sources are projected data based on 1990's or 1980's emissions with the exception of companies located in Atlanta in the NEI 99 database, which have been updated based on a 1999 emission inventory surveyed by Georgia EPD in 2002 (Southwick, 2004).

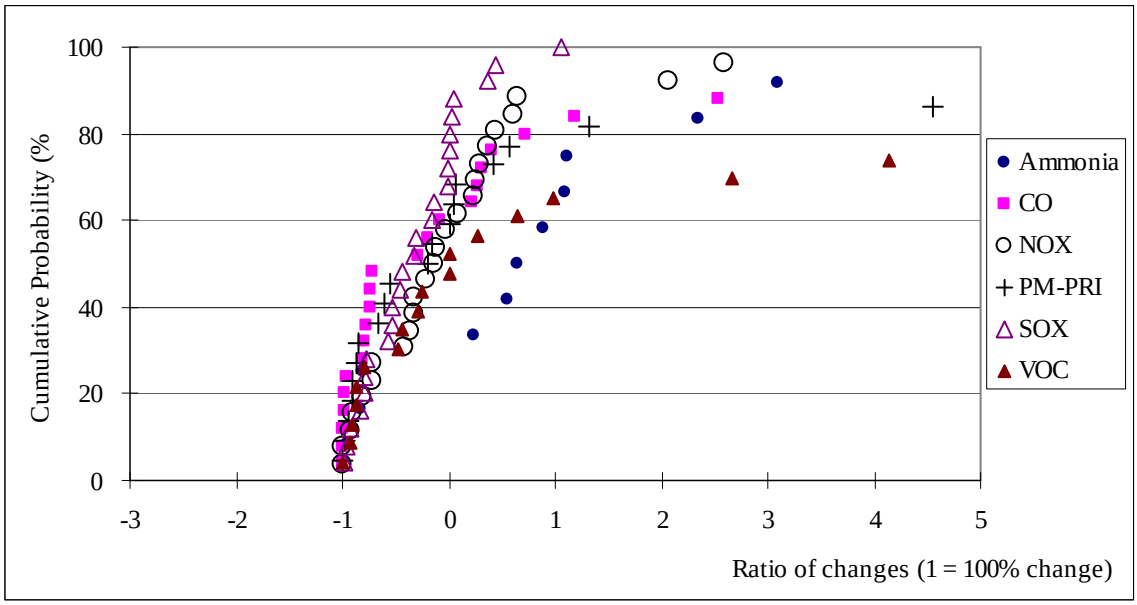


Figure A.4 Cumulative probability distribution of ratios of changes in emissions

In our survey we included 49 companies which do not exist in the NEI 99 database. Table 10 presents emissions for these 49 companies as well as overall emissions from our survey.

Table 10 Increase of emissions by new companies

	NH_3	CO	NO_X	PM10	SO_X	VOC
Emissions from 30 common companies	439	15,304	18,460	3,470	18,265	6,680
Emissions from rest of the 49 companies	8	2,610	5,716	3,419	2,731	3,584
Total emissions	447	17,915	24,176	6,888	20,996	10,264

A.4 CONCLUSION

Efficient and accurate estimation of temporal and spatial variations of emission sources of pollutant precursors is essential in creating a reliable emissions inventory. To acquire the most accurate emissions, a complete survey of every major pollutant source should be implemented. In this study, in order to improve the accuracy of the emission inventory for the state of Georgia (excluding major metropolitan cities), we developed the point source emission inventory for all non-electric generating companies which emit more than 25 tons/year of any of the six pollutants (i.e., CO, SOx, NOx, VOC, PM_{10} and NH_3). Out of 1300 companies 300 of them fell in this category. And of these 300, 79 of them replied to our survey.

We developed software that allowed us to input data into Microsoft Access© database as well as enabling us to do QA/QC on the database. The most frequent errors that were found in QA/QC include using revoked emission factors and reporting combined emission estimate for more than one emission processes. It is found that the percent

change in emissions after correction ranges from 0.1 percent for SO_X, to13 for NH_3 and CO.

Comparison of our emissions to NEI 99 showed that NEI 99 overestimated CO more than four times of surveyed values. For NO_X, PM_{10}, and $SO_{X,}$ this number is between 1.1 and 2.0. For VOCs, NEI 99 underestimated by 1.1 times and it underestimated NH_3 by nine times. Another important finding is the fact that 49 companies that are included in this study do not exist in NEI 99. Total emissions for these 49 companies made up approximately 2,600 tons of CO, 5,700 tons of NO_X, 3,400 tons of PM_{10}, 2,700 tons of SO_X, 3,500 tons of VOC, and 8 tons of NH_3 emissions per year. These two findings indicate that NEI 99 point source inventory for non-egu companies in the state of Georgia (excluding major metropolitan cities) have significant errors. This study successfully identified some of the sources of this uncertainty and developed accurate emissions database.

REFERENCES

Altieri, K. E., Carlton, A. G., Lim, H. J., Turpin, B. J. and Seitzinger, S. P., 2006. Evidence for oligomer formation in clouds: Reactions of isoprene oxidation products. Environmental Science & Technology 40 (16), 4956-4960.

Boylan, J. W., Odman, M. T., Wilkinson, J. G. and Russell, A. G., 2006. Integrated assessment modeling of atmospheric pollutants in the Southern Appalachian Mountains: Part II – PM2.5 and visibility. Journal of Air & Waste Management Association. 56, 12-22.

Butler, A. J., Andrew, M. S. and Russell, A. G., 2003. Daily sampling of PM2.5 in Atlanta: results of the first year of the assessment of spatial aerosol composition in Atlanta study. Journal of Geophysical Research-Atmospheres 108 (D7), 8415.

Butler, A. J., Andrew, M. S. and Russell, A. G., 2003. Daily sampling of pm2.5 in atlanta: Results of the first year of the assessment of spatial aerosol composition in atlanta study. Journal of Geophysical Research-Atmospheres 108 (D7), 8415.

Byun, D. W. and Ching, J. K., 1999. Science algorithms of the epa models-3 community multiscale air quality (cmaq) modeling system, Atmospheric modeling division, National Exposure Research Laboratory, U.S. Environmental Protection Agency.

Carroll, R. J. and Ruppert, D., 1988. Transformation and weighting in regression, Chapman and Hall.

CDC, 2006. State tobacco activities tracking and evaluation (state) system National

Center For Chronic Disease Prevention and Health Promotion.

Cheng, M., Hopke, P. and Jennings, D., 1988. The effects of measurement errors, collinearity and their interactions on aerosol source apportionment computations. Chemmometrics and intelligent laboratory systems 4 (3), 239-250.

Chow, J. and Watson, J., 2002. Review of $PM_{2.5}$ and PM_{10} apportionment for fossil fuel combustion and other sources by the chemical mass balance receptor model. Energy & Fuels 16 (2), 222-260.

Chow, J., Watson, J., Kuhns, H., Etyemezian, V., Lowenthal, D., Crow, D., Kohl, S., Engelbrecht, J. and Green, M., 2004. Source profiles for industrial, mobile, and area sources in the big bend regional aerosol visibility and observational study. Chemosphere 54 (2), 185-208.

Claeys, M., Graham, B., Vas, G., Wang, W., Vermeylen, R., Pashynska, V., Cafmeyer, J., Guyon, P., Andreae, M. O., Artaxo, P. and Maenhaut, W., 2004. Formation of secondary organic aerosols through photooxidation of isoprene. Science 303 (5661), 1173-1176.

CMAS Center, 2007. SMOKE v2.3.2 User's Manual, institute for the Environment - The University of North Carolina at Chapel Hill

Dockery, D. W. and Pope, C. A., 1994. Acute respiratory effects of particulate air-pollution. Annual Review of Public Health 15, 107-132.

Dommen, J., Metzger, A., Duplissy, J., Kalberer, M., Alfarra, M. R., Gascho, A.,

Weingartner, E., Prevot, A. S. H., Verheggen, B. and Baltensperger, U., 2006. Laboratory observation of oligomers in the aerosol from isoprene/NOx photooxidation. Geophysical Research Letters 33 (13), -.

E.H. Pechan & Associates Inc., 2006. Documentation for the final 2002 nonpoint sector (feb. 06 version) National emission inventory for criteria and hazardous air pollutants. Research Triangle Park, U.S. EPA

Edgerton, E. and Jansen, J., 2004. An aries, search, and mercury update. Georgia Environmental Protection Division.

Edgerton, E. S., Hartsell, B. E., Saylor, R. D., Jansen, J. J., Hansen, D. A. and Hidy, G. M., 2005. The southeastern aerosol research and characterization study: Part II. Filter-based measurements of fine and coarse particulate matter mass and composition. Journal of the Air & Waste Management Association 55 (10), 1527-1542

EPA, 1999, National Emission Inventory Documentation and Data

http://www.epa.gov/ttn/chief/NEI/1999inventory.html

EPA, 2000. Background information on the PM Supersites program July 1, 2000, U.S. EPA.

EPA, 2003. 1999 national emission inventory documentation and data - final version 3.0, U.S. Environmental Protection Agency.

EPA, 2004. Emissions modeling: Inventories, U.S. Environmental Protection Agency.

EPA, 2004. Meling platform including emission inventories and smoke ancillary files, U.

S. Environmental Protection Agency.

EPA, 2004. Smoke user manual, U.S. Environmental Protection Agency.

EPA, 2006. Cold temperature effects on vehicle HC emissions. Draft report, Office of Transportation and Air Quality (U.S. Environmental Protection Agency).

EPA, 2006. Speciate 4.0, EPA.

EPA, 2007. AQS Data Dictionary. Office of Air Quality Planning and Standards.

EPRI, 2007. Fact Sheet: SEARCH:: Southeastern Aerosol Research and Characterization.

Evans, D. A., Hobbs, B. F., Oren, C. and Palmer, K. L., 2008. Modeling the effects of changes in new source review on national SO2 and NOx emissions from electricity-generating units. Environmental Science & Technology 42 (2), 347-353.

Fall Line Air Quality Study (FAQS) 2004. FAQS Website.

http://cure.eas.gatech.edu/faqs/emissions_survey/index.html. Located April 27.

Georgia EPD, 2002. The 1999 periodic emissions inventory for the Atlanta, Georgia ozone nonattainment area.

Giannelli, R. A., Gilmore, J. H., Landman, L., Srivastava, S., Beardsley, M., Brzezinski, D., Dolce, G., Koupal, J., Pedelty, J. and Shyu, G., 2002. Sensitivity analysis of mobile6.0. Agency, E. P., Environmental Protection Agency**:** 64.

Gilliland, A. B., Dennis, R. L., Roselle, S. J. and Pierce, T. E., 2003. Seasonal nh3

emission estimates for the eastern united states based on ammonium wet concentrations and an inverse modeling method. Journal of Geophysical Research-Atmospheres 108 (D15)

Griffin, R. J., Cocker, D. R., Flagan, R. C. and Seinfeld, J. H., 1999. Organic aerosol formation from the oxidation of biogenic hydrocarbons. Journal of Geophysical Research-Atmospheres 104 (D3), 3555-3567.

Griffin, R. J., Cocker, D. R., Seinfeld, J. H. and Dabdub, D., 1999. Estimate of global atmospheric organic aerosol from oxidation of biogenic hydrocarbons. Geophysical Research Letters 26 (17), 2721-2724.

Guenther, A., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., Harley, P., Klinger, L., Lerdau, M., Mckay, W. A., Pierce, T., Scholes, B., Steinbrecher, R., Tallamraju, R., Taylor, J. and Zimmerman, P., 1995. A Global-Model of Natural Volatile Organic-Compound Emissions. Journal of Geophysical Research-Atmospheres 100 (D5), 8873-8892.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I. and Geron, C., 2006. Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). Atmospheric Chemistry and Physics 6, 3181-3210.

Hansen, D. A., E. S., Hartsell, B. E., Jansen, J. J., Kandasamy, N., Hidy, G. M. and Blanchard, C. L., 2003. The southeastern aerosol research and characterization study: Part 1 – overview. Journal of Air & Waste Management Association. 53,

1460-1471.

Hedberg, E., Johansson, C., Johansson, L., Swietlicki, E. and Brorstrom-Lunden, E., 2006. Is levoglucosan a suitable quantitative tracer for wood burning? Comparison with receptor modeling on trace elements in Lycksele, Sweden. Journal of the Air & Waste Management Association 56, 1669-1678.

Held, T., Ying, Q., Kleeman, M., Schauer, J. and Fraser, M., 2005. A comparison of the UCD/CIT air quality model and the CMB source-receptor model for primary airborne particulate matter. Atmospheric Environment 39 (12), 2281-2297.

Hildemann, L. M., Markowski, G. R. and Cass, G. R., 1991. Chemical-composition of emissions from urban sources of fine organic aerosol. Environmental Science & Technology 25 (4), 744-759.

Hopke, P. K., Ramadan, Z., Paatero, P., Norris, G. A., Landis, M. S., Williams, R. W. and Lewis, C. W., 2003. Receptor modeling of ambient and personal exposure samples: 1998 Baltimore Particulate Matter Epidemiology-Exposure Study. Atmospheric Environment 37 (23), 3289-3302.

Hopke, P. K., Ito, K., Mar, T., Christensen, W. F., Eatough, D. J., Henry, R. C., Kim, E., Laden, F., Lall, R., Larson, T. V., Liu, H., Neas, L., Pinto, J., Stolzel, M., Suh, H., Paatero, P. and Thurston, G. D., 2006. PM source apportionment and health effects: 1. Intercomparison of source apportionment results. Journal of Exposure Science and Environmental Epidemiology 16 (3), 275-286.

Iinuma, Y., Boge, O., Gnauk, T. and Herrmann, H., 2004. Aerosol-chamber study of the

alpha-pinene/O-3 reaction: influence of particle acidity on aerosol yields and products. Atmospheric Environment 38 (5), 761-773.

IMPROVE, 1995. IMPROVE data guide, University of California, Davis.

Jang, M. S., Czoschke, N. M., Lee, S. and Kamens, R. M., 2002. Heterogeneous atmospheric aerosol production by acid-catalyzed particle-phase reactions. Science 298 (5594), 814-817.

Jang, M., Czoschke, N. M., Northcross, A. L., Cao, G. and Shaof, D., 2006. SOA formation from partitioning and heterogeneous reactions: Model study in the presence of inorganic species. Environmental Science & Technology 40 (9), 3013-3022.

Kalberer, M., 2004. Identification of polymers as major components of atmospheric organic aerosol. Science 303, 1659-1662.

Kim, E., Hopke, P. K., Paatero, P. and Edgerton, E. S., 2003. Incorporation of parametric factors into multilinear receptor model studies of Atlanta aerosol. Atmospheric Environment 37 (36), 5009-5021.

Lane, T. E., Pinder, R. W., Shrivastava, M., Robinson, A. L. and Pandis, S. N., 2007. Source contributions to primary organic aerosol: Comparison of the results of a source-resolved model and the chemical mass balance approach. Atmospheric Environment 41 (18), 3758-3776.

Lee, S., Kim, H. K., Yan, B., Cobb, C. E., Hennigan, C., Nichols, S., Chamber, M.,

Edgerton, E. S., Jansen, J. J., Hu, Y. T., Zheng, M., Weber, R. J. and Russell, A. G., 2008. Diagnosis of aged prescribed burning plumes impacting an urban area. Environmental Science & Technology 42 (5), 1438-1444.

Lee, S., Liu, W., Wang, Y., Russell, A. and Edgerton, E. S., 2008. Source apportionment of $PM_{2.5}$: Comparing pmf and cmb results for four ambient monitoriniz sites in the southeastern united states. Atmospheric Environment 42 (18), 4126-4137.

Lemire, K. R., Allen, D. T., Klouda, G. A. and Lewis, C. W., 2002. Fine particulate matter source attribution for southeast texas using c-14/c-13 ratios. Journal of Geophysical Research-Atmospheres 107 (D22)

Lewis, C. W., Klouda, G. A. and Ellenson, W. D., 2004. Radiocarbon measurement of the biogenic contribution to summertime pm-2.5 ambient aerosol in nashville, tn. Atmospheric Environment 38 (35), 6053-6061.

Liu, C., Roscoe, B., Severin, K. and Hopke, P., 1982. The application of factor-analysis to source apportionment of aerosol mass. American industrial hygiene association journal 43 (5), 314-318.

McInnes, G. EMEP/CORINAIR Emission Inventory Guidebook – 3rd edition. European Environment Agency, Technical reports no 30. 2002.

MACTEC Inc., 2005. Documentation of the revised 2002 base year, revised 2018, and initial 2009 emission inventories for vistas, Visibility Improvement State and Tribal Association of the Southeast (VISTAS).

Marmur, A., Park, S. K., Mulholland, J. A., Tolbert, P. E. and Russell, A. G., 2006. Source apportionment of $PM_{2.5}$ in the southeastern United States using receptor and emissions-based models: Conceptual differences and implications for time-series health studies. Atmospheric Environment 40 (14), 2533-2551.

Marmur, A., Unal, A., Mulholland, J. and Russell, A., 2005. Optimization-based source apportionment of $PM_{2.5}$ incorporating gas-to-particle ratios. Environmental Science & Technology 39 (9), 3245-3254.

McDonald, J., Zielinska, B., Fujita, E., Sagebiel, J., Chow, J. and Watson, J., 2003. Emissions from charbroiling and grilling of chicken and beef. Journal of Air & Waste Management Association 53 (2), 185-194.

Mendoza, A. 2001. A four-dimensional data assimilation method for air quality modeling. School of Civil and Environmental Engineering. Atlanta, Georgia Institute of Technology.

Mysliwiec, M. and Kleeman, M., 2002. Source apportionment of secondary airborne particulate matter in a polluted atmospbere. Environmental Science & Technology 36 (24), 5376-5384.

Napelenok, S. L., Carlton, A. G., Bhave, P. V., Sarwar, G., Pouliot, G., Pinder, R. W., Edney, E. O. and Houyoux, M., 2008. Updates to the treatment of secondary organic aerosol in CMAQ v4.7. the 7th Annual CMAS Conference. Chapel Hill, NC.

NARSTO, 2004. Particulate matter science for policy makers. P. McMurry, M. Shepherd

and J. Vickery.

National Research Council, 2004. Air Quality Management in the United States National Research Council. Committee on Air Quality Management in the United States.

Ng, N. L., Kroll, J. H., Keywood, M. D., Bahreini, R., Varutbangkul, V., Flagan, R. C., Seinfeld, J. H., Lee, A. and Goldstein, A. H., 2006. Contribution of first- versus second-generation products to secondary organic aerosols formed in the oxidation of biogenic hydrocarbons. Environmental Science & Technology 40 (7), 2283-2297.

Paatero, P. and Hopke, P., 2002. Utilizing wind direction and wind speed as independent variables in multilinear receptor modeling studies. Chemometrics and Intelligent Laboratory Systems 60 (1-2), 25-41.

Paatero, P., Hopke, P. K., Hoppenstock, J. and Eberly, S. I., 2003. Advanced factor analysis of spatial distributions of pm2.5 in the eastern united states. Environmental Science & Technology 37 (11), 2460-2476.

Pace, T. G., 2005. Methodology to estimate the transportable fraction (tf) of fugitive dust emissions for regional and urban scale air quality analyses. EPA.

Park, S.-K., 2005. Particulate modeling and control strategy for atlanta, georgia. School of Civil and Environmental Engineering. Atlanta, Georgia Institute of Technology.

PECHAN, 2004. Summary of approaches available for fugitive dust sources. Roe, S. M. and Hemmer, P., Mid-Atlantic Regional Air Management Association**:** 38.

Peel, J. and Tolbert, P., 2002. Ambient air pollution and respiratory emergency department visits in Atlanta, August 1998-August 2000 (ARIES/SOPHIA). Epidemiology 13 (4), S124-S124.

Peel, J. L., Tolbert, P. E., Klein, M., Metzger, K. B., Flanders, W. D., Todd, K., Mulholland, J. A., Ryan, P. B. and Frumkin, H., 2005. Ambient air pollution and respiratory emergency department visits. Epidemiology 16 (2), 164-174.

PSU/NCAR, 2003. Psu/ncar mesoscale modeling system tutorial class notes and user's guide: Mm5 modeling system version 3. Mesoscale and microscale meteorology division, national center for atmospheric research.

Pun, B. K., Seigneur, C., Vijayaraghavan, K., Wu, S. Y., Chen, S. Y., Knipping, E. M. and Kumar, N., 2006. Modeling regional haze in the BRAVO study using CMAQ-MADRID: 1. Model evaluation. Journal of Geophysical Research-Atmospheres 111 (D6), D06302.

Robinson, A. L., Donahue, N. M. and Rogge, W. F., 2006. Photochemical oxidation and changes in molecular composition of organic aerosol in the regional context. Journal of Geophysical Research-Atmospheres 111 (D3), D03302.

Robinson, A. L., Donahue, N. M., Shrivastava, M. K., Weitkamp, E. A., Sage, A. M., Grieshop, A. P., Lane, T. E., Pierce, J. R. and Pandis, S. N., 2007. Rethinking organic aerosols: Semivolatile emissions and photochemical aging. Science 315 (5816), 1259-1262.

Roe, S. M. and Hemmer, P., 2004. Summary of approaches available for fugitive dust

sources, Technical Memorandum. Mid-Atlantic Regional Air Management Association. Baltimore, MD.

Roelle, P. A., Aneja, V. P., Gay, B., Geron, C. and Pierce, T., 2001. Biogenic nitric oxide emissions from cropland soils. Atmospheric Environment 35 (1), 115-124.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R. and Simoneit, B. R. T., 1998. Sources of fine organic aerosol 9. Pine, oak and synthetic log combustion in residential fireplaces. Environmental Science & Technology 32 (1), 13-22.

Sakulyanontvittaya, T., Duhl, T., Wiedinmyer, C., Helmig, D., Matsunaga, S., Potosnak, M., Milford, J. and Guenther, A., 2008. Monoterpene and sesquiterpene emission estimates for the United States. Environmental Science & Technology 42 (5), 1623-1629.

Sakulyanontvittaya, T., Guenther, A., Helmig, D., Milford, J. and Wiedinmyer, C., 2008. Secondary Organic Aerosol from Sesquiterpene and Monoterpene Emissions in the United States. Environmental Science & Technology 42 (23), 8784-8790.

Sarnat, J. A., Marmur, A., Klein, M., Kim, E., Russell, A. G., Sarnat, S. E., Mulholland, J. A., Hopke, P. K. and Tolbert, P. E., 2008. Fine particle sources and cardiorespiratory morbidity: An application of chemical mass balance and factor analytical source-apportionment methods. Environmental Health Perspectives 116 (4), 459-466.

Schauer, J. and Cass, G., 2000. Source apportionment of wintertime gas-phase and particle-phase air pollutants using organic compounds as tracers. Environmental

Science & Technology 34 (9), 1821-1832.

Schauer, J. J., ., Rogge, W. F., Hildemann, L. M., Mazurek, M. A. and Cass, G. R., 1996. Source apportionment of airborne particulate matter using organic compounds as tracers. Atmospheric Environment 30 (22), 3837-3855.

Schauer, J., 2003. Evaluation of elemental carbon as a marker for diesel particulate matter. Journal of Exposure Analysis and Environmental Epidemiology 13 (6), 443-453.

Schauer, J., Kleeman, M., Cass, G. and Simoneit, B., 2001. Measurement of emissions from air pollution sources 3. C-1-C-29 organic compounds from fireplace combustion of wood. Environmental Science & Technology 35 (9), 1716-1728.

Schlesinger, R., Kunzli, N., Hidy, G., Gotschi, T. and Jerrett, M., 2006. The health relevance of ambient particulate matter characteristics: Coherence of toxicological and epidemiological inferences. Toxicology. 8 (2), 95-125.

Schnelle-Kreis, J., Sklorz, M., Peters, A., Cyrys, J. and Zimmermann, R., 2005. Analysis of particle-associated semi-volatile aromatic and aliphatic hydrocarbons in urban particulate matter on a daily basis. Atmospheric Environment 39, 7702-7714.

Scott Southwick, 2004. Georgia Environmental Protection Division of the Georgia Department of Natural Resources, *personal communication*.

Seinfeld, J. H. and Pandis, S. N., 2006. Atmospheric chemistry and physics: From air pollution to climate change. 2nd edition. Wiley Interscience.

Sheesley, R. J. and Schauer, J. J., 2004. Source apportionment of atmospheric fine

particulate matter collected at the Seney national wildlife refuge. Environmental Chemistry and Technology Program. LADCO. Des Plaines, IL.

Shu, Y. H. and Atkinson, R., 1995. Kinetics of the Gas-Phase Reactions of a Series of Sesquiterpenes with Oh Radicals, No3 Radicals and O3. Abstracts of Papers of the American Chemical Society 209, 108-ENVR.

Simoneit, B. R. T., Rogge, W. F., Lang, Q. and Jaffe, R., 2000. Molecular characterization of smoke from campfire burning of pine wood (Pinus elliottii). Chemosphere - Global Change Science 2, 107-122.

Tian, D., Hu, Y. T., Wang, Y. H., Boylan, J. W., Zheng, M. and Russell, A. G., 2009. Assessment of Biomass Burning Emissions and Their Impacts on Urban and Regional PM2.5: A Georgia Case Study. Environmental Science & Technology 43 (2), 299-305.

Tie, X. X., Li, G. H., Ying, Z. M., Guenther, A. and Madronich, S., 2006. Biogenic emissions of isoprenoids and no in china and comparison to anthropogenic emissions. Science of the Total Environment 371 (1-3), 238-251.

Turpin, B. J., Saxena, P. and Andrews, E., 2000. Measuring and simulating particulate organics in the atmosphere: problems and prospects. Atmospheric Environment 34, 2983-3013.

United States Department of Agriculture, 2006. Food availability on-line data. (http://www.ers.usda.gov/data/foodconsumption

Van Donkelaar, A., Martin, R. V., Park, R. J., Heald, C. L., Fu, T. M., Liao, H. and Guenther, A., 2007. Model evidence for a significant source of secondary organic aerosol from isoprene. Atmospheric Environment 41 (6), 1267-1274.

Watson, J. G., Cooper, J. A. and Huntzicker, J. J., 1984. The effective variance weighting for least-squares calculations applied to the mass balance receptor model. Atmospheric Environment 18 (7), 1347-1355.

Yarwood, G., Morris, R. E. and Wilson, G. M., 2004. Particulate matter source apportionment technology (PSAT) in the CAMx photochemical grid model. The International technical Meeting, Banff, Canada.

Ying, Q., Mysliwiec, M. and Kleeman, M., 2004. Source apportionment of visibility impairment using a three-dimensional source-oriented air quality model. Environmental Science & Technology 38 (4), 1089-1101.

Zhang, K. M. and Wexler, A. S., 2004. Evolution of particle number distribution near roadways - Part I: analysis of aerosol dynamics and its implications for engine emission measurement. Atmospheric Environment 38 (38), 6643-6653.

Zhang, K. M., Wexler, A. S., Niemeier, D. A., Zhu, Y. F., Hinds, W. C. and Sioutas, C., 2005. Evolution of particle number distribution near roadways. Part III: Traffic analysis and on-road size resolved particulate emission factors. Atmospheric Environment 39 (22), 4155-4166.

Zhang, Y., Huang, J. P., Henze, D. K. and Seinfeld, J. H., 2007. Role of isoprene in secondary organic aerosol formation on a regional scale. Journal of Geophysical

Research-Atmospheres 112 (D20), D20207.

Zheng, M., Cass, G., Schauer, J. and Edgerton, E., 2002. Source apportionment of $PM_{2.5}$ in the southeastern United States using solvent-extractable organic compounds as tracers. Environmental Science & Technology 36 (11), 2361-2371.

Zheng, M., Ke, L., Edgerton, E. S., Schauer, J. J., Dong, M. Y. and Russell, A. G., 2006. Spatial distribution of carbonaceous aerosol in the southeastern United States using molecular markers and carbon isotope data. Journal of Geophysical Research-Atmospheres 111 (D10), D10S06.

Zielinska, B., Sagebiel, J., McDonald, J. D., Whitney, K. and Lawson, D. R., 2004. Emission rates and comparative chemical composition from selected in-use diesel and gasoline-fueled vehicles. Journal of Air & Waste Management Association. 54, 1138-1150.